Review

Modeling and optimization of environment in agricultural greenhouses for improving cleaner and sustainable crop production

The corrections made in this section will be reviewed and approved by a journal production editor.

Yu Guo[a,b,c], Huajian Zhao[a,b,c], Shanhong Zhang[a,b,c], Yang Wang[a,bc,*] andy_yangwang@cau.edu.cn, David Chow[d]

[a]College of Information and Electrical Engineering, China Agricultural University, 100083, Beijing, China

[b]National Innovation Center for Digital Fishery, China Agricultural University, 100083, Beijing, China

[c]Beijing Engineering and Technology Research Center for Internet of Things in Agriculture, China Agricultural University, 100083, Beijing, China

[d]School of Architecture, University of Liverpool, L69 3BX, Liverpool, UK

*Corresponding author. College of Information and Electrical Engineering, China Agricultural University, 100083, Beijing, China., (

Handling editor: Prof. Jiri Jaromir Klemeš

Abstract

Resource-use efficiency and crop yield are significant factors in the management of agricultural greenhouse. Appropriate modeling methods effectively improve the control performance and efficiency of the greenhouse system and are conducive to the design of water and energy-saving strategies. Meanwhile, the extreme environment could be forecasted in advance, which reduces pests and diseases as well as provides high-quality food. Accordingly, the interest of the scientific community in greenhouse modeling and optimizing has grown considerably. The objective of this work is to provide guidance and insight into the topic by reviewing 73 representative articles and to further support cleaner and sustainable crop production. Compared to the existing literature review, this work details the approaches to improve the greenhouse model in the aspects of parameter identification, structure and process optimization, and multi-model integration to better model complex greenhouse system. Furthermore, a statistical study has been carried out to summarize popular technology and future trends. It was found that dynamic and neural network techniques are most commonly used to establish the greenhouse model and the heuristic algorithm is popular to improve the accuracy and generalization ability of the model. Notably, deep learning, the combination of "knowledge" and "data", and coupling between the greenhouse system elements have been considered as future valuable development.

Keywords: Agricultural greenhouse; Environment; Modeling and optimization; System identification; Heuristic algorithm; Multi-model integration

1 Introduction

All along, sustainability and land as well as resource requirements for food have attracted the major concern from agriculturists and society (Lagerberg and Brown, 1999). However, in the past few decades, the reckless exploitation of natural resources by human beings has caused a global effect on the ecosystem (Aiello et al., 2018). The global population could grow to around 9.7 billion in 2050 (United Nations et al., 2019). This means that the demand for food and other production is continuing to rise rapidly, and the existing natural resources are also under increasing pressure (Mirzamohammadi et al., 2020). FAO (2018) indicted that high-external input, resource-intensive agricultural systems have caused massive deforestation, water scarcities, biodiversity loss, soil depletion, and high levels of greenhouse gas emissions. 70% of mankind's water consumption and 13% of greenhouse gas emissions are occupied by agriculture (Fox et al., 2019). Thus, efficiently increasing resource-use efficiency and recycling of resources are critical and urgent issues (FAO, 2018). In the situation, agroecology embodies great value as a sustainable food and agriculture system.

Where, agricultural greenhouse, one of the Controlled Environment Agriculture (CEA), is a powerful assistant for the development of the agroecology.

According to the latest statistics, there are an estimated 3.64 million hectares of greenhouses worldwide (McNulty, 2017). Agricultural greenhouse provides an ideal environment for crop growth and food production and allows crops to be produced in the seasons that would otherwise inhibit growth, which prolongs the cultivation period of seasonal crops and raises the production (Kamal, 2013). While ensuring production safety, the demand for continuous food has been achieved throughout the year (Baddadi et al., 2019). Furthermore, a suitable greenhouse environment effectively prevents pests and disease, promotes the crop metabolic activities, as well as reduces water and land consumption (Iddio et al., 2020; M.C. M.C. Singh et al., 2018). This merit directly reduces the external inputs of fertilizer and pesticide, which further decrease environmental pollution and carbon emissions, and provides cleaner and high-quality food production. In addition to crop farming, the greenhouse system is also used in fields of crop drying, aquaculture, soil solarization, and poultry (Choab et al., 2019; Khanlari et al., 2020).

The design and management of greenhouse is a multi-factor optimization problem, which affects crop productivity and the resource-use efficiency of land, water, and energy. An efficient greenhouse system is essential for sustainable production. For solving this issue, Vanthoor et al. (2011) proposed a model-based method to simulate the greenhouse productive procession as climate model, crop yield model, and economic model. Agricultural greenhouse system modeling could deepen the understanding of the relationship between controlled environment and crop production.

Specifically, the necessity of greenhouse modeling comes from the following reasons: **1)** The control for greenhouse environment is the key in successful greenhouse system operation, and this environment control accounts for 65 to 85 percent of greenhouse energy demand (Fox et al., 2019). Because of many user-adjustable control settings and complex environment-crop interactions, optimal control of greenhouse system is a difficult task (Llera et al., 2019). Determined model is beneficial to design the controller and understand the mechanism in the executive process of the greenhouse system (Iddio et al., 2020; López-Cruz et al., 2018; Wang and Wang, 2020). In order to implement advanced environmental control (e.g., adaptive, feedback, intelligent and model predictive control) and make better decision-making, it is necessary to build a greenhouse model to predict the process of mass and energy change accurately (J. Li et al., 2017; Lin et al., 2020). **2)** In addition, one problem mentioned by Li et al. (2002) is that high temperatures and untimely ventilation or too low temperatures and untimely heating will harm crop growth. Seriously, this untimely control actions will cause widespread crop death and bring substantial financial loss to farmers. While, the greenhouse climate model can predict the trend of environmental factors, which helps farmers anticipate extreme environment and produce better crops (Elanchezhian et al., 2020). **3)** Furthermore, designers could use the model for greenhouse structure optimization under the characteristics of different regions (Esmaeli and Roshandel, 2020; Vanthoor et al., 2011a, 2011b; Zhao and Wang, 2010).

Because the effectiveness of interaction between crop and environment as well as the utilization ration of resources and energy are the interests of sustainable agroecology, the discussed greenhouse model was explained as mathematical or physical description of the relationship with greenhouse environments (i.e. climates or microclimates), crops, and actuators. According to the model, the relationship between control strategies of actuators and optimal environments for maintaining crop growth is explored. Furthermore, the changes of energy consumption with microclimate conditions and other factors could be directly captured (Grabarczyk, 2018). For this reason, the greenhouse model could provide the basis for the design of the control strategy with low energy consumption (Ahamed et al., 2019; Laktionov et al., 2020; Liang et al., 2018; Zhang et al., 2020).

Due to the aforementioned contributions of the greenhouse system model (i.e. efficient and accurate environmental control, forecasting and early warning, and low energy consumption management) to ensure cleaner and sustainable crop production, a large number of studies addressing the greenhouse modeling strategies have been conducted in the recent years. The objective of this paper is to provide a systematic and holistically review to respond to the following research questions: What are the advanced greenhouse modeling methods? How to optimize and improve the model to enhance performance? What are the trends and potential inadequacies of the current greenhouse model?

2 Literature review

In this section, the existing literature reviews on greenhouse modeling and related fields were analyzed, and an attempt has been made to figure out the areas that were not been given due attention.

In the past few decades, a larger number of mechanism model which emphasized physical processes were researched to increase the knowledge of system and control the greenhouse (López-Cruz et al., 2018). Therefore, several literature summarized the process and experience of mechanism modeling and simulation. Sethi et al. (2013) reviewed thermal modeling methods for greenhouse microclimate control, and have categorized the thermal model to three categories (i.e. independent model, heating dependent model, and cooling dependent model). Singh et al. (2016) indicated that greenhouse climate is the major driving force that directly affects crop production. Hence, greenhouse microclimate parameters, models (static and dynamic), and application were summarized to better understand the relationships between microclimate and plants community. In addition to describing the mathematical model, Choab et al. (2019) reviewed the application of Computational Fluid Dynamics (CFD) simulation in greenhouse thermal modeling.

Cunha (2003) devoted an overview of greenhouse microclimate modeling approaches, such as physical-based models, black-box linear parametric models, and black-box nonlinear parametric models. López-Cruz et al. (2018) demonstrated that mechanistic and black-box model were benefited to optimize and control greenhouse system, and presented that uncertainty analysis and sensitivity analysis are required for dynamic models in order to increase their reliability. On the basis of previous literature, Iddio et al. (2020) described two advanced simulation methods (i.e. system dynamic and adaptive machine learning approaches), and simply summarized the hybrid model optimized for the state-space model. Because data-oriented black-box model could fit the complex system which cannot be solved by physical-based model, Escamilla-García et al. (2020) reviewed the application of artificial neural networks in greenhouse technology.

The existing literature reviews provide a good starting point for researchers but also have the following limitations:

- The current advanced methods for greenhouse modeling (e.g. machine learning and deep learning) were rarely described, systematically.
- The single models have been paid too much attention, and the potential values of hybrid models were ignored.
- The focus point was on the modeling approach without an in-depth overview of the possibilities of improvement and optimization.

Because of the characteristics of strong coupling, nonlinearity, and strong disturbance, the agricultural greenhouse is extremely complex and is not necessarily well modeled with simple approaches, especially when high precisions are required (Li et al., 2016; Lin et al., 2019; Xu et al., 2019). Therefore, several optimization approaches have been encouraged for enhancing the performance of greenhouse model. In this respect, it is evident from the literature that no comprehensive review on the scheme of improving and optimizing greenhouse model has been published so far. For the purpose of it, we have conducted an in-depth review of the advanced greenhouse modeling technologies in recent years. Besides, the different optimization and reinforcement methods have been holistically explored to help establish the model with higher accuracy and generalized performance for complex agricultural greenhouse system.

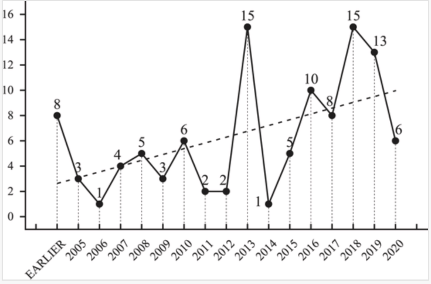
3 Method

An intensive literature review was conducted from September 2019 to April 2020 in order to explore efficient and more robust methods for modeling the greenhouse system. The review collected current relevant works derived from the following databases: *Web of Science, ScienceDirect, Scopus,* and *Google Scholar*, which allow accurate and customized searches. The main keywords for the review were selected to determine the suitable scientific literature, including five categories: “agricultural greenhouse”, “system modeling”, “parameter identification”, “optimization”, and “model integration”. During the retrieval process, we iteratively expanded the search strategy of keywords according to the reviewed literature information. The final keywords were combined according to the following search string [(“agricultural greenhouse” OR “greenhouse”) AND (“system modeling” OR “model” OR “simulation”) AND (“system identification” OR “parameter identification” OR “optimization” OR “heuristic algorithm” OR “particle swarm optimization” OR “genetic algorithm” OR “integration” OR “hybrid” OR “ensemble” OR “prior knowledge”)], where Boolean search terms, namely AND OR, were used to incorporate diverse but reasonable keywords in one search string.

Specifically, a preliminary search on the bibliographic database using a set of simple keywords, and obtained 3951 articles after deduplication. Then, further filtering was made based on the title and content of abstracts. After that, full-text was analyzed to select out 55 articles, and several rules for a more polished search were defined. Finally, a refined search was carried out by the optimal rules, and 18 articles were expanded in the collection. In doing so, 73 of the most representative articles were reviewed in detail and depth. In conjunction with the supporting materials, 123 articles were cited in this work. According to the time period distribution of the primary literature up to April 2020, it is evident that research on accurate and efficient modeling for agricultural greenhouse has been on the rise (see Fig. 1). In the last six years (i.e., from 2015 to 2020), there is a considerable number of researches focused on the topic of greenhouse modeling.

alt-text: Fig. 1

Fig. 1

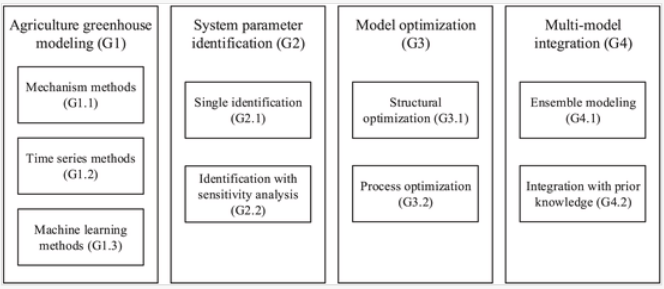


Distribution of the 73 main articles across the time period.

Moreover, the classification for works of 73 main articles was executed by research method, applied model, and optimization technique for modeling agricultural greenhouses. Then, we analyzed the relationship of each article and archived them to distinct and homogeneous groups. As depicted in Fig. 2, the literature was classified into four groups: the group 1 (G1) comprehensively describes general and original methods for modeling greenhouse; the group 2 (G2) includes applications for identifying uncertain parameters of greenhouse mechanism model; the group 3 (G3) covers optimized algorithms for greenhouse black-box model; and the group 4 (G4) expounds the ideas of multi-model integration. Furthermore, the four groups were divided into several sub-groups according to the detail of the technique. The organization of the “Results” section was designed by the above groups, and several figures were drawn to explain the work and aim of each article. Finally, we established a table to summarize the information of reviewed literature (Table A.1 in “Appendix A”). In the table, the influence factors for modeling were divided into five aspects according to the summary by Su and Xu (2017). **Outside climate factors**: outside air temperature, humidity, solar radiation, wind speed, and CO_2 concentration; **Inside climate factors**: Inside air temperature, humidity, soil temperature, illumination, and CO_2 concentration; **Actuators for control**: CO_2 injecting, ventilation, heating, and fogging. **Crop behaviors**: photosynthesis rate, transpiration rate, respiration rate, leaf area index, dry matter of crop, and fruit. **Dimensional and physical parameters**: plant density, height, and floor area. (see Fig. 3).

alt-text: Fig. 2

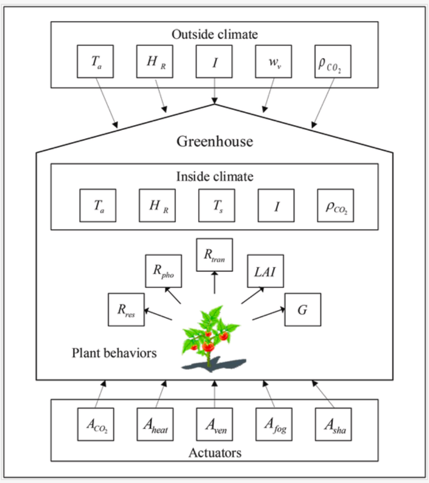
Fig. 2



Classification of research focuses from 73 main articles.

alt-text: Fig. 3

Fig. 3



Influence factors for modeling agricultural greenhouse system (T_a is air temperature; H_R is relative humidity; I is Solar radiation; w_v is wind speed; ρ_{co_2} is CO_2 concentration; T_s is soil temperature; R_{res} is respiration rate; R_{pho} is photosynthesis rate; R_{tran} is transpiration rate; LAI is leaf area index; G is growth stage; A_{CO_2} is CO_2 injecting; A_{heat} is heating; A_{ven} is ventilation; A_{fog} is fogging; A_{sha} is shading).

4 Results

In this section, current greenhouse modeling methods were described in the first sub-section with the aims of reviewing classic methods, introducing advanced techniques, and pointing out the merit and dilemma of different models. This is the basis for model optimization and helps to understand the reason and purpose of the improvement. Then, the improvement methods were introduced, detailedly, in the remaining three sub-sections based on the dilemma of the existing model.

4.1 Agricultural greenhouse modeling

According to the literature review, this sub-section is divided into three categories by different technologies for greenhouse modeling: mechanism methods, time series methods, and machine learning methods. For the above category, the model by the latter two methods is referred to as the Greenhouse Black-box Model (GBM) in this paper, which means that these methods do not have to pay attention to the law and principles of physics in the greenhouse. In addition to the above three kinds of methods, the fuzzy theory, Peteri Nets, and other technologies are applied to greenhouse modeling by some researchers (Salgado and Cunha, 2005; Tovany et al., 2013; Trabelsi et al., 2007; Yaofeng et al., 2018).

4.1.1 Mechanism methods

Greenhouse Mechanism Model (GMM) uses the law of physiological and physics principles to analyze related factors in the greenhouse quantitatively. It is based on energy conservation and mass conservation principle to establish a balance equation. GMM can be decomposed into the static model and dynamic model. Static models are variously called steady-state models, as opposed to the dynamic model that reveals how a variable develops in time (Rhinehart, 2016). Dynamic models are typically described by differential equations, and it describes the change rule of the system.

The earliest static model of the greenhouse was established by Businger (1963), which laid a foundation for later research. Although the static model is easy to implement, its accuracy is low. Wang and Boulard (2000) indicated that the usefulness of static models decreases when the time response of the greenhouse becomes comparable with the rates of time change of the boundary conditions. Based on those reasons, Takakura et al. (1971) built the first relatively complete dynamic model for the unheated symmetrical span greenhouse. This model comprehensively described the heat and moisture transfer process in the greenhouse. Roni Avissar and Ytzhaq Mahrer (1982) developed a greenhouse microclimate model, which comprised of balance equations of soil layer, crops layer, air layer, and cover layer. The change of temperature in the greenhouse was simulated by convective exchange coefficient, heat flux, and outdoor solar radiation.

Since 1970, a number of similar models have been proposed by researchers around the world. Except for some differences in greenhouse structure, system assumptions, parameter determination, and solution methods, the basic ideas of modeling are consistent. Namely, based on fully understanding the mathematical expressions of the physical processes, the executive process of the greenhouse is revealed by solving the greenhouse mass and energy balance equations for each system component subjected to the interference and initial conditions. Cunha (2003) summarized the greenhouse dynamic model into the following general formulation.

$$\begin{aligned} \frac{dT_i}{dt} &= \frac{1}{C}\left(q_{in,h} - q_{out,h} + p_h\right) \\ \frac{dc_m}{dt} &= \frac{1}{V}\left(q_{in,m} - q_{out,m} + p_m\right) \end{aligned} \tag{1}$$

where, T_i is the inside air temperature, C is the thermal capacity, $q_{in.h}$ and $q_{out,h}$ are the energy (heat) inflow and outflow, and p_h means the energy (heat) production per unit time. Likewise, C_m is the mass concentration, $q_{in.m}$ and $q_{out,m}$ represent the mass inflow and outflow, and p_m is the mass production per unit time referred to the greenhouse volume $V\left(m^3\right)$.

In recent years, the greenhouse dynamic model has been improved and developed. One of the previous research, a greenhouse microclimate model under cucumber crop in soilless media was developed to describe the energy and mass transport processes (Mahesh Chand Singh et al., 2018). This model was capable of predicting the temperature of air, plant, growing media, and plastic cover under natural ventilation. Joudi and Farhan (2015) and Mohammadi et al. (2018) used the dynamic heat transfer model and experimentally validated the performance of the inside environment model in an innovative greenhouse structure. The innovation greenhouse can effectively reduce the load and cost of greenhouse cooling and heating.

To compensate for the defect that lumped parameters cannot calculate the greenhouse temperature and humidity distribution (spatial characteristics), a new greenhouse simulation method, Computational Fluid Dynamics (CFD), has been widely used (Choab et al., 2019). Zhou et al. (2017) used the CFD software Airpak to model the greenhouse climate. Firstly, the continuous space was divided into several subspaces through the unstructured grid generation, and the calculation domain of the entire simulation model was constituted by a finite number of discrete points. On this basis, the unsteady calculation capacity of Fluent was used to simulate the processes of inside space temperature and wind speed field in natural and mechanical ventilation, and the effectiveness of the model was verified. Finally, through the obtained model, the optimal climate control scheme for crop growth was searched, which was based on multi-objective optimization algorithm via three aspects of energy consumption, CO_2 concentration, and air temperature. Saberian and Sajadiye (2019) used CFD to study dynamic solar heat load and temperature field under the influence of radiation and natural convection inside the greenhouse. The result of performance of the numerical model proved that CFD methods could predict variable solar heat load and inside temperature in several hours of a day. The requirement of solar removal and ventilation during hot months was effectively suggested by these methods. Zhang et al. (2016) established a computer simulation model by CFD for evaluating the temperature distribution of Chinese solar greenhouses in winter nights. It is demonstrated that a desirable thickness of the north wall can improve energy conservation. Therefore, weighted entropy and fuzzy optimization methods were employed to achieve the best north wall thickness. Since the characteristics of CFD, such as high-order and time-consuming, perform better in analysis, rather than control, Li et al. (2020) brought Proper Orthogonal Decomposition (POD) into the environment parameter description. Three subspaces of low dimensionality parameters were constructed by POD, and then the subspaces were combined with a multi-objective evolution algorithm. Moreover, CFD data was used to find the optimal values of environment parameters within the crop area, which reduced the calculation cost and increased the real-time performance while ensuring special high resolution.

For previous studies of GMM, a large number of articles took the greenhouse structure, materials, and thermal environment components as main research objects. Most studies considered less impact of crop physiological activity for greenhouse and the model subjected to extremely stringent assumptions. Hence, it is a challenge to generalize GMM in practice. Besides, GMM uses a lot of physical variables and parameters, which will cause a considerable complexity for modeling.

4.1.2 Time-series methods

According to the greenhouse system features of significant delay and slow time variation, the time series data of greenhouse environment factors have a specific sequence variation trend and periodic characteristics. Many researchers have used Greenhouse Time Series Model (GTSM) to explore the specific rule in the environment series data. From the perspective of the greenhouse system, the time series at a particular time represents a dynamic process of the objective system. Further, a time sequence can be regarded as the relevant output or response of the system.

4.1.2.1 Conventional statistic methods

Conventional time series analysis methods establish a linear series model (Autoregressive Exogenous (ARX), Autoregressive Moving Average Exogenous (ARMAX)) or nonlinear series model (Nonlinear ARX (NARX)) through the rule of Auto-Correlation Function (ACF) and Partial Auto-Correlation Function (PACF) analyzed by statistical methods to estimate the target factor. Uchida Frausto et al. (2003) pointed out that climate characteristics are continuous variables, but they are measured and registered at time steps, which give measured climate data and a discrete character. Therefore, linear autoregressive can be used to model the dynamic greenhouse system. The ARX and ARMAX models were used to fit the synthetic data generated by a validated simulation greenhouse model. When the time span of input data is more than 15 min, this data did not improve the model performance. From the result, ARX performs better than ARMAX except for the ventilation periods. In order to solve the problem of slow time-variant identification for the greenhouse ARX model, Coelho et al. (2005) brought a Recursive Least Squares Algorithm with Exponential Forgetting (RLSEF) into the identification and designed a greenhouse air temperature model predictive controller. Hui et al. (2017) described a nonlinear greenhouse temperature system approximated by the ARMAX model and analyzed the correlation of input and output variables. Then, fading memory recursive least squares was adopted to identify parameters of ARMAX. The result showed that this model had excellent versatility and adaptability.

4.1.2.2 Deep learning methods

Traditional autoregressive methods only focus on changes in the time series itself, lack the ability to mine the information of exogenous series, and cannot recognize the structure and pattern of nonlinear or complex systems. In recent years, several deep learning methods have been widely developed in multivariate time series analysis tasks (Liu et al, 2019, 2020; Qin et al., 2017). Recurrent Neural Network (RNN) has a context layer, which is used to save the output state of the hidden layer at the current moment, to represent the historical characteristics of the object (Wang et al., 2018). Hence, it can process the data of sequence changes compared with the classical neural network. Wang et al. (2018) modified the weights of the RNN by Back Propagation (BP) algorithm and used the RNN model to predict future air temperature and humidity through multivariate time series data. Moreover, Salah and Fourati (2020) built a Deep Elman RNN model by climate and actuator variables and stated that the obtained model would be used to the control task.

For the long sequence prediction task, RNN has trouble with vanishing or exploding gradient. Hochreiter and Schmidhuber (1997) proposed Long Short-Term Memory (LSTM) model, which adds a memory structure to each base component of RNN. It is clear that this property dramatically eases the restriction of learning the temporal dependency compared to RNN. Kim et al. (2018) pointed out that LSTM showed better performance of forecasting in the non-stationary environment and long-term time lags. Moon et al. (2018) used LSTM to predict future Electrical Conductivity (EC) of nutrient solutions of crops and claimed that LSTM could predict crop environment affected by the accumulations of historical situations. To avoid adverse effects of high or low temperatures on crops, Ali and Hassanein (2019) and Ángel et al. (2019) forecasted in advance whether or not the extreme temperature inside the greenhouse will occur by LSTM model. Furthermore, Jung et al. (2020) respectively compared performance among with Artificial Neural Network (ANN), NARX, and LSTM for time series task of temperature, humidity, CO_2 in greenhouse. The result represented that LSTM not only had the highest overall accuracy, but also held least impact of accuracy decreases as prediction time increases.

4.1.3 Machine learning methods

With the development of Artificial Intelligence, the calculated performance and capacity of data production were further improved, which promotes the Greenhouse Machine Learning Model (GMLM) to be applied in greenhouse modeling more and more.

4.1.3.1 Artificial neural network methods

Artificial Neural Network (ANN) is a bionic intelligent information processing method for simulating the human brain neural system. It has a robust comprehensive information processing ability and could handle irregular and nonlinear multi-parameter data. Whereas its ability to process the small sample is weak, the convergence speed is slow, and it is easy to fall into the local optimum.

Taki et al. (2016) compared Dynamic, Multiple Linear Regression (MLR), and ANN methods for predicting energy loss. The result of the *t*-test, F-test, and Kolmogorov-Smirnov-test for the ANN model indicated that the predicted data series equal with the actual data series. It was found that the dynamic model has terrible results to calculate inside roof temperature by Least Significant Difference method. Trejo-Perea et al. (2009) used an ANN of cascading architecture to predict energy consumption, where the outputs of temperature and relative humidity model were used as inputs for the predictor. The result of Duncan's Multiple Range Test indicated that ANN was approximately similar to real data with a 95% confidence level compared to the regression model. Yue et al. (2018) proposed a model to predict the temperature and humidity of a greenhouse based on improved Levenberg-Marquardt Radial Basis Function Neural Network (LM-RBFNN), and this model basically realizes the prediction.

In order to improve the generalization performance and training speed of the neural network, He and Ma (2010) extracted the four main factors from eight environmental factors influencing the inside humidity by Principal Component Analysis (PCA), and used it for the input of Back Propagation Neural Network (BPNN). As a result, BPNN based on PCA was found to be significantly superior to the stepwise regression model. Zou et al. (2015) proposed an improved Extreme Learning Machine (ELM) based on Orthonormal Basis Function (OBF) to accelerate training speed. Compared to the traditional ELM, the prediction results of the improved ELM showed that temperature error and humidity errors are reduced by 2 and 5%, respectively. Liu et al. (2016) claimed the Kernel-based ELM (KELM) model requires less training time but showed more energetic fitness and more stable performance, compared with the prediction models based on BP, Elman, and SVM.

4.1.3.2 Support vector machine methods

Support Vector Machine (SVM) based on the principle of minimized structural risk can be processed to the dataset of the small sample. The input sample space can be mapped to a high dimensional linear feature space by applying the nonlinear kernel function. Consequently, SVM can do highly nonlinear classification and regression.

The greenhouse environment is an uncertain nonlinear system. Therefore, Wang et al. (2009) claimed that SVM offers a very competent method for modeling the greenhouse system. In order to solve the uncertainty, the model was rectified online, and the real-world data was used to prove the usefulness of Online Sparse Least-Squares SVM (OS-LSSVM). Yu et al. (2016) proposed a temperature prediction model for typical solar greenhouse based on LSSVM. The prediction model can provide forecasts for temperature within a time interval of 6 h, and provide accurate forecasts within 4 h. Yang and Zhao (2012) successfully used LSSVM in the wind prevailing direction forecasting for natural ventilation of the greenhouse. For their test, the prediction accuracy was satisfying in 10 min and better than ANN model. It was mentioned that the wind direction prediction, compared with wind speed, is difficult. In the research of Yan et al. (2016), it was regarded as binary data whether the greenhouse environment is beneficial to plant growth, and agricultural experts were invited to calibrate it. In order to solve the problem of data imbalance, the original proximal SVM model was enhanced, and the weighted proximal SVM method was proposed to optimize the decision-making system for the environmental monitoring of greenhouse crop growth.

Given all above-mentioned review, a prediction SVM model with high accuracy can be obtained by balancing the small sample information and the error threshold design. SVM model overcomes shortcomings of the traditional neural network, but it is prone to local extremum and calculation disaster of the high dimensional sample.

Machine learning methods adequately use statistical analysis and train the model through the samples. These methods can effectively obtain a model with high accuracy and small number of manual operation parameters for complex nonlinear greenhouse system. However, these methods never have the capability of detailed geometric modeling. The model is built according to the data. Therefore, they cannot apply to the conceptual model and fail to establish in the situation of being a lack of data. Because the greenhouse system is time-varying, it is difficult to tune in practice, and can only provide reasonable predictions for the short-term future time range.

4.2 System parameter identification

The advantage of mechanism methods is that the influence rule of parameters for the greenhouse system can be realized, such as the structural characteristics of the greenhouse, the inside climatic environment, the control equipment, and the growth conditions of crops by GMM. In other words, to establish the GMM, researches should have sufficient expertise and ability to grasp the details of greenhouse material and structure, energy and mass executive process, and crop behavior (Choab et al., 2019). Obviously, for a simple greenhouse system, it is possible. With the development of recent decades, the greenhouse application has risen dramatically, the greenhouse structure has been continuously updated to complex, and the number of elements to be considered has been increasing. The outside environment and the crops grown in greenhouses are uncertain with the change of market and season, and the mechanism models of transpiration and photosynthesis of crops are inaccurate (López-Cruz et al., 2018). It is challenging to build a complete GMM only based on the experience of researchers. Therefore, more and more researchers have applied data-driven methods in the optimization of GMM, which not only reduces the complexity of model construction but also reflects the understanding of the mechanism of the greenhouse system. Where, System Parameter Identification (SPI) is a typical representative technology. In the mechanism modeling of the greenhouse with a simple structure, the parameter values of each part of the model are often obtained through experience, manual or reference materials. However, for the complex greenhouse system, researchers may not adequate cognized the specific value of model parameters, and only determined a setting range of Uncertain-Parameters (UPs). For this situation, SPI can be used to assist the establishment of GMM, which will obtain an equivalent model that conforms to the law of the real system.

4.2.1 Technology of parameter identification

System identification is a theory to establish a practical model for the industrial production process by using input and output data. According to some criteria, System identification can provide a basis for designing control systems and improving the quality of control methods. For instance, within the range of the model cluster, a model equivalent to the identified system was determined (Zadeh, 1962). Ljung (1978) explained that system identification has three elements: data, model cluster, and criteria. Among them, data is the basis of identification; Criteria is regarded as optimization objective; And, the model cluster is scope for searching a model. The essence of identification is to select a model from a set of the model cluster and to make it better fit the dynamic characteristics of the actual system process according to specific criteria.

At present, the linear system identification theory has become mature, and yet the general linear model is an approximate mathematical description of the real system after certain nonlinear is ignored or replaced with a linear relationship. The greenhouse is a typical nonlinear system, cannot be simulated by the linear mathematics model (Blasco et al., 2007). Therefore, the traditional linear identification methods (e.g., Least-Squares, Maximum Likelihood) is not well-suited to GMM. With the development of identification science, new methods such as neural network and swarm intelligence have been proposed to identify the complex system. Furthermore, these methods have made some achievements in the identification of the complex and nonlinear systems.

SPI is a subset of system identification, which is a process of determining the UPs of the model according to input and output data. It could be well used in GMM. There are many methods of parameter identification, where, the heuristic algorithm has apparent advantages, such as fast convergence speed and excellent global search performance. It is highly praised by researchers solving some non-differentiable, discrete, and nonlinear problems.

4.2.2 Abstract representation of greenhouse model

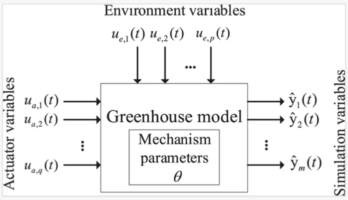
Through the analysis of GMM based on the law of heat and mass transfer and energy conservation, the parameters and variables related with GMM were divided into the following three classes: 1) Environment variables, also known as disturbances variable, include climate, and crop behaviors (Hasni et al., 2010); 2) Actuator variables are viewed as the execution state of controlled device, such as heating, cooling, and fogging; 3) Mechanism parameters contain greenhouse material parameter, structure information, and heat transfer coefficient. It could be divided into time-varying and non-time-varying parameters. SPI could solve those mechanism parameters which cannot be determined by experience. For the convenience of discussion, the greenhouse dynamic model can be adapted directly to a state-space model Blasco et al. (2007); Chen et al. (2016); Hasni et al. (2011)0.

$$\begin{aligned} \dot{\boldsymbol{x}}(t_k) &= f(\boldsymbol{x}(t_k), \boldsymbol{u}(t_k); \boldsymbol{\theta}) \\ \hat{\boldsymbol{y}}(t_k) &= g(\boldsymbol{x}(t_k), \boldsymbol{u}(t_k); \boldsymbol{\theta}) \end{aligned} \tag{2}$$

where, $f(\cdot)$ and $g(\cdot)$ are the function to represent greenhouse system; t_k is used to mark k-th sample, and researchers usually conduct sampling at a fixed time interval T; $\boldsymbol{\theta}$ is uncertain parameters vector; $\boldsymbol{u}$ is the input vector, which includes p-dimension environment variables $\boldsymbol{u}_e$ and q-dimension actuator status variables $\boldsymbol{u}_a$; $\boldsymbol{x}$ is state vector, which explains the transformation of greenhouse dynamic model and needs initialization; and $\hat{\boldsymbol{y}}$ is the m-dimension output/simulation vector of greenhouse model, namely, the predicted targets of the model. Thus, the GMM is abstracted to the block diagram shown in Fig. 4.

alt-text: Fig. 4

Fig. 4



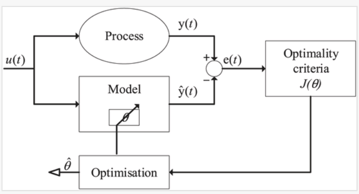
The abstract block diagram of the greenhouse model.

4.2.3 Application of parameter identification

After determining the GMM and their UPs, the SPI can be carried out by designing of data sampling. Hasni et al. (2011, 2010) optimized the value of UPs and natural ventilation and provided a proper schema for the application of SPI in greenhouse dynamic model. Yang et al. (2019) used the Increased Convergence Factor Particle Swarm Optimization (ICFPSO) algorithm to optimize the eight UPs of greenhouse temperature and humidity model. On this basis, considering the disturbance from random environmental effects, the stochastic dynamic model was built by the maximum likelihood estimation method. As described by Herrero et al. (2007), the SPI techniques for GMM can be summarized as follows: First, the research defined the search space for UPs, and updated the value of UPs in this space by the optimization algorithm. Then, the output of the greenhouse model was calculated by the input sample $u(t)$ of environment and actuators state for every new parameter. Finally, in the algorithm iteration, the optimal UPs $\hat{\theta}$ was found, which was equal to the parameter θ used in the system with minimum error $e(t)$ between the model simulating output $\hat{y}(t)$ and the actual measurement $y(t)$, where the error is evaluated by optimality criteria $J(\theta)$ (see Fig. 5).

alt-text: Fig. 5

Fig. 5



The general procedure of parameter identification by Herrero et al. (2007) ($u(t)$ is model input; $y(t)$ is process output; $\hat{y}(t)$ is model output; $e(t)$ is identification error and $\hat{\theta}$ is optimal parameters).

In the UPs of the greenhouse model, a part of parameters is affected by the environment variables, while others are affected by actuator variables, and there is no natural coupling between meteorological conditions and controlled equipment status. Nonetheless, the environment and actuator variables are commonly bound together and used to construct datasets in sample design of SPI. In order to simplify SPI process of GMM, Cruz-Valeriano et al. (2013) divided the optimization process into two stages: For the first stage, the parameters related to the characteristic of system were only identified; And for the second stage, based on the previous stage, the parameters associated with status of actuators were identified. Finally, the Particle Swarm Optimization (PSO) algorithm was used to obtain better precision.

The greenhouse is strictly a multi-output system. For example, inside temperature, vapor pressure, and CO_2 concentration are taken as the considered targets of system (Lammari et al., 2012). One method to solve the problem of multi-output model identification is to handle the difference between the simulating and measurement output vector in a particular time t_k by L^2-norm, and the cost function for optimizing is as follows:

$$J(\boldsymbol{\theta}) = \frac{1}{m \times N} \sum_{k}^{N} \left\| \boldsymbol{y}(t_k) - \hat{\boldsymbol{y}}(t_k, \boldsymbol{\theta}) \right\|_2^2 \tag{3}$$

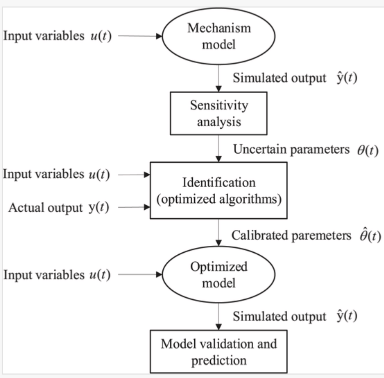
where, N is the number of samples; m is the dimension of output vector; $\boldsymbol{\theta}$ is uncertain parameters to be optimized; $\boldsymbol{y}(t_k)$ is the actual measurement vector; $\hat{\boldsymbol{y}}(t_k, \boldsymbol{\theta})$ is the model simulating output vector.

There is often a strong coupling among vector elements of the greenhouse model simulating output, and the approach of Eq. (3) does never well balance the importance and relationship among vector elements. Herrero et al. (2007) proposed a Multi-Objective Evolutionary Algorithm (MOEA) for identifying greenhouse model with hydroponic roses. The optimality criteria were to minimize the third quartile of the identification error of the two outputs in the model. The algorithm independently designed evaluation metrics for inside temperature and humidity. It avoided having to make a priori decision about the relative importance of the fit of the inside humidity and temperature. On this basis, a method of robust identification was mentioned (Herrero et al., 2008).

As the GMM becomes complex, the number of UPs in the model increases. Because the calculative complexity of the optimization algorithm is raised exponentially with the increasing of the dimension of search space, the identification cost of the algorithm will increase dramatically. Chen et al. (2016) effectively reduced the number of UPs in the dynamic model of complex greenhouse through sensitivity analysis based on Sobol' method. Before performing SPI, researchers calculated the first-order sensitivity index and total sensitivity index of 12 UPs through Sensitivity Analysis (SA). Then, three parameters of third-class were excluded and assigned with a constant. Finally, nine parameters to be identified were obtained (see Fig. 6). Based on the SA, it is found that the parameters with high sensitivity have a high correlation with others, which also revealed the strongly coupled characteristics of the greenhouse system. Besides, the sum of all first-order indices is less than 1, which suggests that the model is non-additive. In their study, a hybrid model Adaptive PSO and Genetic Algorithms (APSO-GA) was used to identify the UPs of the system. The results showed that the accuracy and convergence speed of the algorithm are superior to the PSO and GA (Table 1), and these methods could benefit to combine the advantage of a large number of GA and larger inertia weight in PSO. Shen et al. (2018) also reduced the dimension of search space by SA. Based on the final optimization model, the energy consumption was reduced in the day with extreme weather through reasonable changing temperature distribution in a week according to accumulated temperature theory. Guzmán-Cruz et al. (2009) found that, through SA, the parameter associated with opening of windows and evapotranspiration of the crop was very sensitive to the temperature and humidity model of greenhouse, and sixteen parameters to be identified were finally determined. By comparing three global evolutionary algorithms (GAs, Evolution Strategy (ES) and Evolution Programming (EP)) and two local search methods (Least Squares (LSQ) and Sequential Quadratic Programming (SQP)), seen from Table 2, it was found that the model obtained by EP algorithm has better performance for temperature and humidity. On the contrary, the estimation of relative humidity by LSQ and SQP is very inaccurate, which may be affected by the local minimum.

alt-text: Fig. 6

Fig. 6



The flowchart of parameter identification with sensitivity analysis from Chen et al. (2016).

alt-text: Table 1

Table 1

The table layout displayed in this section is not how it will appear in the final version. The representation below is solely purposed for providing corrections to the table. To preview the actual presentation of the table, please view the Proof.

The performance of four algorithms for optimizing uncertain parameters in Chen et al. (2016).

Evaluation indicator	PSO	GA	CPSO-GA	APSO-GA
Exit generation	76	66	58	50
Time(s)	832.5	2525.4	758.2	664.2
RMSE	187,654	187,382	187,281	187,277

(Note: RMSE is root mean square error).

alt-text: Table 2

Table 2

The table layout displayed in this section is not how it will appear in the final version. The representation below is solely purposed for providing corrections to the table. To preview the actual presentation of the table, please view the Proof.

The performance of global evolutionary and local search algorithms in Guzmán-Cruz et al. (2009).

Method	Air temperature				Air humidity			
	r	E	%SEP	ARV	r	E	%SEP	ARV
No optimization	0.9181	0.3597	23.2265	0.6403	0.9254	0.6669	15.0927	0.3331
GAs	0.9183	0.8426	11.5159	0.1574	0.9075	0.4747	18.9539	0.5253
EP	0.9187	0.8429	11.5046	0.1571	0.9332	0.7989	11.728	0.2011
ES	0.9185	0.8433	11.4896	0.1567	0.9263	0.5414	17.711	0.4586
LSQ	0.9183	0.7804	13.6016	0.2196	0.9303	0.3451	21.1637	0.6549
SQP	0.9187	0.844	11.4655	0.156	0.9328	0.4288	19.7648	0.5712

(Note: r is correlation coefficient; E is efficiency coefficient; %SEP is percentage standard error of the prediction; ARV is average relative variance).

For many previous studies, researchers have treated the UPs of GMM as time-invariant parameters. However, in the actual implementation, most parameters of the model are changing with real-time (e.g., different season, morning and evening), because of the influence of various complex factors, such as natural or mechanical ventilation and crop growth state. In order to obtain a more accurate model, online identification technology is needed (Pérez-González et al, 2014, 2018). Firstly, Pérez-González et al. (2018) explored the optimal calibration parameter and performance of the optimization algorithm by offline identification method (see Table 3). According to comparison, the Classical PSO algorithm had premature convergence and often obtained a locally optimal solution; Many Optimizing Liaisons PSO (MOLPSO) restrained the phenomenon of premature convergence by increment population number; ICFPSO also did never converge prematurely, but its accuracy was better than that of MOLPSO. On the premise of ensuring the global optimum, offline-PSO algorithms spent much time. The execution time of the optimal ICFPSO algorithm was 152 h 24 m, with 86,400 samples. Differential Evolution (DE) current to best/1 algorithm reduced the execution time to 18 h 06 m under the accuracy that MSE of model was 9.5040, and yet the DE algorithm is not significantly improved in the online situation. The optimal uncertain parameter obtained in the offline phase was used to generate the initial particles population of online identification task. In order to reduce the execution time of the algorithm, namely making it less than 1s sampling time, researchers obtained a new evaluation function (sample to sample):

$$J(\theta;k) = \frac{1}{m}\left\|y(t_k) - \hat{y}(t_k;\theta)\right\|_2^2 \quad (4)$$

where, m is the dimension of output vector; k is used to mark the current time point; θ is uncertain parameters to be optimized. This function evaluates the performance through the sum of the residual squares by the model output of sample at a single time point under the particle. Under the task of online, ICFPSO can effectively track the time-varying trajectory of UPs, which was significantly improved by MSE = 5.64.

alt-text: Table 3

Table 3

The table layout displayed in this section is not how it will appear in the final version. The representation below is solely purposed for providing corrections to the table. To preview the actual presentation of the table, please view the Proof.

The performance of PSOs and DEs algorithms for offline identification in Pérez-González et al. (2018).

Method	MSE	Total execution time
Classical PSO	17.9975	153 h 24 m
MOLPSO	11.3021	172 h 28 m 55s
ICFPSO	10.8761	152 h 24 m
DE rand/1	10.8806	17 h 53 m 19.8s
DE current to best/1	9.5040	18 h 06 m 19.8s

(Note: MSE is mean square error).

4.3 Structural and process optimization

Due to the characteristics of MIMO, strong coupling, large inertia, and nonlinearity in the greenhouse system, the original machine learning method has some disadvantages, such as local optimization and slow convergence. Many researchers have used heuristic algorithms to optimize the structure and learning process of GBM, namely GTSM or GMLM, to overcome the shortcomings of original machine learning and further to improve the model accuracy and generalization performance. **Structural Optimization (SO)**, namely model selection, mainly uses an optimization algorithm to select the hidden layer node of neural network or hyperparameters of SVM, which improves the randomness and blindness of parameter setting for GBM. **Process Optimization (PO)** brings disturbance into the training process of GBM and improves the iteration speed of the algorithm and performance of global search. At present, the hybrid method of machine learning and intelligent optimization algorithm is a popular topic in greenhouse system modeling. Many studies demonstrated that this kind of optimization algorithm could effectively improve the prediction performance of GBM.

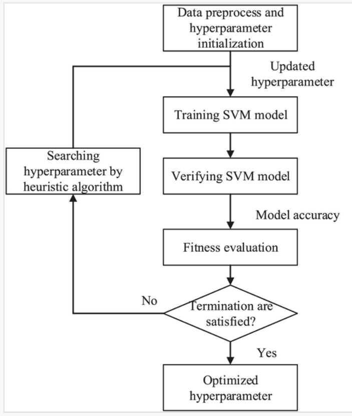
4.3.1 Optimization for SVM

For the construction of SVM model in the greenhouse system, the selection of hyperparameters has a significant influence on the performance of the final model (Bao et al., 2013; Huang and Dun, 2008; Huang and Wang, 2006). The following several methods are often used to determine suitable hyperparameters: cross-validation, grid search, random search, Nelder-Mead search, a heuristic search, pattern search, etc. (Steinwart and Christmann, 2008). When the number of models hyperparameter is more than two, the parameter selection method based on the exhaustive search (e.g., grid search) becomes intractable (Chapelle et al., 2002). In the meanwhile, as the search grid density increases, the time cost also rises sharply. The heuristic algorithm has an excellent sophisticated continuous space search capability, as well as advantages of fast, simple, and easy-programming. More and more researchers have used heuristic methods to select hyperparameters of SVM.

García Nieto et al. (2016) developed a PSO-RBF-SVM model for Chlorophyll-a of Spirulina platensis in greenhouses by eight input data associated with greenhouse climate and experimental open raceway ponds. For the experiment, the hyperparameter of three different kernel function SVM models by PSO algorithm was selected. The results indicated that PSO algorithm could effectively improve the generalization ability of the model, and coefficients of determination (R^2) of model was 0.99. Jian et al. (2018) respectively optimized the hyperparameter of SVM model for photosynthetic rate prediction through grid search method and PSO algorithm and founded that the accuracy of PSO was higher than grid search. Given all above literature, the process of hyperparameter optimization for SVM is summarized as Fig. 7.

alt-text: Fig. 7

Fig. 7



The flowchart for optimizing hyperparameters of SVM by heuristic algorithm form Yu et al. (2016).

Since the Classical PSO algorithm cannot guarantee the convergence of the model to the optimal value, namely it is prone to local optimal solution, and has the shortcomings of high complexity. Most researchers have improved the application of the Classical PSO algorithm. Yu et al. (2016) created a mutation probability p_m into PSO, which improved global convergence of algorithm. If the variance of population? was small and no particles converged to the theoretical optimal value, each particle should be multiplied by a random value in the range of [1,2] under the probability p_m. The result demonstrated that Classical PSO is easy to fall into local minimum and slow convergence speed; The convergence speed of DE is fastest; But, Improved PSO (IPSO) can run closer to the fitness evolutions while accelerating the speed. Also, the effect of temperature prediction was compared from two perspectives: horizontal (BP, SVM and IPSO-SVM) and vertical (time scale: 1 h, 2 h, 4 h, and 6 h). As shown in Table 4, the comparison illustrated that the performance of IPSO-SVM among three models is the best, and this method provides accurate temperature predictions for time interval up to 4 h. Li et al. (2017b, 2017a) improved the convergence speed and global search performance of PSO from the aspect of nonlinear processing of learning factors and inertia weight. The photosynthesis of the three growth stages of tomato through IPSO-SVM model was forecasted, and coefficients of determination (R^2) were increased by 0.07, 0.09, and 0.08, respectively, compared with the SVM model optimized by Classical PSO (see Table 5). Then, with the aid of this model, the effect of CO_2 enrichment on tomato photosynthesis was studied.

alt-text: Table 4

Table 4

The table layout displayed in this section is not how it will appear in the final version. The representation below is solely purposed for providing corrections to the table. To preview the actual presentation of the table, please view the Proof.

The performance of different methods for modeling greenhouse temperature in Yu et al. (2016).

Model	MAE	MAPE	MSE
BPNN	0.8765	0.1415	1.0113
SVM	0.8115	0.1354	0.8747
IPSO-SVM	0.1062	0.0279	0.0281

(Note: MAE is mean absolute error; MAPE is mean absolute percentage error; MSE is mean squared error).

alt-text: Table 5

Table 5

The table layout displayed in this section is not how it will appear in the final version. The representation below is solely purposed for providing corrections to the table. To preview the actual presentation of the table, please view the Proof.

The performance of two models for various growth stages in Li et al. (2017b, 2017a).

Evaluation indicator	PSO-SVM			IPSO-SVM		
	LSS	FFS	EFS	LSS	FFS	EFS
R^2	0.89	0.87	0.86	0.96	0.96	0.94
ARE	31	33	29	11	9	12

RMSE	2.73	2.87	3.30	1.34	1.35	1.96
MAR	2.20	2.36	2.60	1.09	1.20	1.58

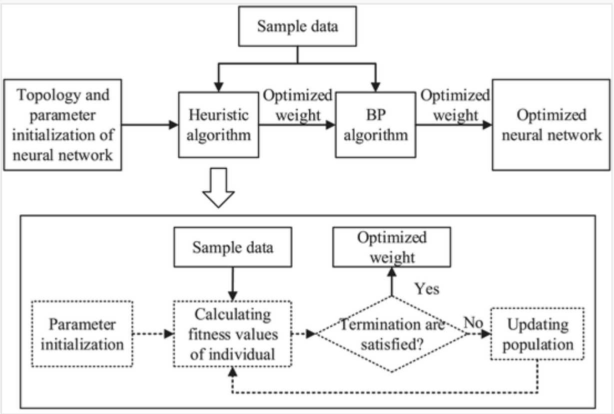
(Note: LSS is late seeding stage; FFS is full flowering stage; EFS is early fruiting stage; R^2 is determination coefficient; ARE is average relative error; MAE is mean absolute error; RMSE is root mean square error).

4.3.2 Optimization for ANN and time series model

The BP algorithm based on gradient descent is easy to make the neural network of prediction for greenhouse environment trap to a local optimum, and the convergence rate is slow. Outanoute et al. (2018) used PSO algorithm to search optimal weights and biases parameters of neural network. The result demonstrated that the convergence of the PSO-ANN was very fast, and the excellent learning efficiency and generalization ability were shown compared to the standard neural network algorithm. Although the heuristic algorithm has an excellent global search ability, it does not make sufficient use of the structural characteristics of neural network. The two-stage method integrating heuristic and BP algorithm is used to train the network. This method is shown in Fig. 8: Firstly, the heuristic algorithm was used to find the initial weights and thresholds; Then, BP algorithm was applied to train the network further. The global search ability of heuristic algorithm and the local search ability of BP algorithm were richly utilized. In order to better exploit the advantages of the two-stage method, He et al. (2007) designed the Improved BP (IBP) algorithm by adding the inertia impulse and self-adaptation learning rate. The RMSE of GA-IBP neural network between temperature predicted and measured was 0.8 °C, and the relative humidity RMSE was 1.1%. Guo et al. (2019) balanced the global search and local search ability of algorithm based on nonlinear processing for inertia weights of PSO algorithm. Seen from Fig. 9, the nonlinear PSO-BP network has smallest relative error. So, the demand of air humidity in different growth periods of crops could be timely controlled by this model.

alt-text: Fig. 8

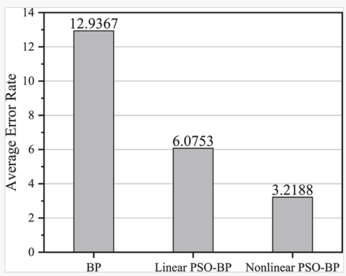
Fig. 8



The process for optimizing ANN by hybrid Heuristic-BP algorithm to describe works of He et al. (2007) and Guo et al. (2019).

alt-text: Fig. 9

Fig. 9



The performance of three algorithms to optimizing neural network in Guo et al. (2019).

Radial Basis Function Neural Network (RBFNN), compared with the BPNN, can approximate any nonlinear mapping and has the characteristics of straightforward and fast convergence speed. Due to the over-fitting phenomenon of RBF neural network to noise data, some researchers have enhanced the generalization performance of model by Regularized Orthogonal Least Squares (ROLS) algorithm (Chen et al., 1996). However, the Bayesian method used to determine regression coefficient λ will make the network fall into local optimization, for the reason that the function constituted by RBFNN generalization with spread σ and regularized coefficients λ is the complex multipeak space surface. Zhao and Wang (2010) conducted a global search on λ, σ by PSO. Compared with optimizing network topology directly, this method required less computation. As shown in Table 6, RBFNN model based on PSO is simpler than the model based on the Orthogonal Least Squares (OLS) algorithm. In the meanwhile, the performance of greenhouse temperature and humidity enhanced significantly.

alt-text: Table 6

Table 6

The table layout displayed in this section is not how it will appear in the final version. The representation below is solely purposed for providing corrections to the table. To preview the actual presentation of the table, please view the Proof.

The performance of RBFNN based on PSO and OLS algorithm in Zhao and Wang (2010).

Method	Temperature		Relative humidity		Number of hidden units
	Mean error	RMSE	Mean error	RMSE	
RBF-OLS	−0.0141	0.2803	0.0013	2.2103	83
RBF-PSO	0.0061	0.2407	−0.0009	2.0692	57

(Note: RMSE is root mean square error).

To solve the local optimization defects of least squares estimation method, Guzmán-Cruz et al. (2013) estimated coefficients in both ARX and ARMAX models for predicting the behavior of air temperature inside a greenhouse, using two evolutionary algorithms (GAs and EP). For the implementation, the greenhouse climate data is divided into five groups for training and testing. The better fit of coefficients of ARX model with the data group 20%:80% for air temperature of greenhouse was performed by GAs (see Table 7). Since this group used 80% of the data for generalization performance testing, there was enough identification and validation. The result of this collocation was very reliable. Moreover, RBFNN was used as the input function of ARX for endowing the model with nonlinear property (Ferreira and Ruano, 2008). The number of neurons, the input variable, and lagged input terms for each variable of network were optimized by Multi-Objective GA (MOGA). This algorithm helped model to obtain several stronger performances, such as minimizing complexity, good generalization ability, and smallest possible long-term prediction errors. Finally, four interconnected models were hybrid, and a better long-term predicting effect was obtained.

alt-text: Table 7

Table 7

The table layout displayed in this section is not how it will appear in the final version. The representation below is solely purposed for providing corrections to the table. To preview the actual presentation of the table, please view the Proof.

The performance of different algorithm for optimizing ARX and ARMAX model with the data group of 20:80% from Guzmán-Cruz et al. (2013).

Method	ARX				ARMAX			
	r	E	%SEP	AVR	r	E	%SEP	AVR
Typical	0.2345	−71.9648	377.0408	72.9648	0.5286	−3.0448	88.7727	4.0448
GAs	0.9296	0.8335	18.0094	0.1665	0.7468	0.5537	29.4889	0.4463
EP	0.9184	0.8428	17.5015	0.1572	0.914	0.8317	18.1064	0.1683

(Note: r is correlation coefficient; E is efficiency coefficient; %SEP is percentage standard error of the prediction; ARV is average relative variance).

4.4 Multi-model integration

A lot of engineering problems, especially environmental and energy modeling, are too complicated for a single model, and the ensemble method is an effective technique. The shortcomings of a single greenhouse model are solved by combining multiple models with apparent diversity, fusing advantage of these models to achieve better results.

4.4.1 Ensemble modeling

With the purpose of improving the performance of heat energy consumption model, Jovanović et al. (2015) used the two-stage design process to integrate three different network architectures (Feed Forward backpropagation Neural

Network (FFNN), RBFNN, Adaptive Neuro-Fuzzy Interference System (ANFIS)): The researchers first generated individual network which was usually trained independently of each other. Then, these networks were integrated by three different combining methods (simple, average, and weighted), and the performance of the integrated model were compared in Table 8. It was found that the ensemble model could effectively integrate the advantages of a single neural network. Esen et al. (2015) integrated neural network (13 neurons for hidden), SVM (RBF kernel), and k-Nearest Neighbor (KNN) (the k value as 3) to model the BSGSHP system. The results of the ensemble model were compared with single neural network by three-fold cross-validation, as well as the ensemble model increased by 1.33% on R^2, decreased by 0.1306 on RMS and reduced by 5.2147 on COV. The structure of above ensemble model could be abstracted as Fig. 10.

alt-text: Table 8

Table 8

The table layout displayed in this section is not how it will appear in the final version. The representation below is solely purposed for providing corrections to the table. To preview the actual presentation of the table, please view the Proof.

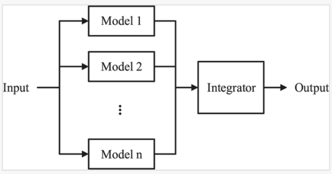
The performance of three individual models and ensemble model with different integrated methods by all available input variables in Jovanović et al. (2015).

Model	R^2	RMSE (kWh)	MAPE (%)
FFNN	0.9814	8496.1	5.6283
RBFNN	0.9816	8849.1	5.6682
ANFIS	0.9783	9115.4	5.5778
Ensemble			
Simple	0.9845	8169.1	5.3204
Weighted	0.9843	8153.6	5.3686
Median	0.9830	8451.4	5.4820

(Note: R^2 is coefficient of determination; RMSE is root mean square error; MAPE is mean absolute percentage error).

alt-text: Fig. 10

Fig. 10

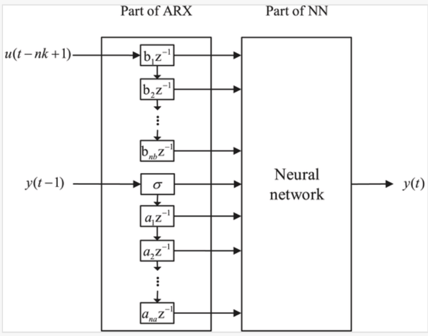


Stacking ensemble model structure from Jovanović et al. (2015).

General ANN model never has feedback and delay, which makes the static neural network ignore the time sequence relationship of the variable itself and leads the model to fail to sufficiently utilize the characteristics of large inertia of the greenhouse. Therefore, researchers have begun to combine ARX and ANN model (Neural Network ARX, NNARX) to model the greenhouse system, which enabled the neural network to equip with the capacity of memory and more candidate signals, and the sequence relationship between input and output could be well-reflected. NNARX model could be abstracted as Fig. 11.

alt-text: Fig. 11

Fig. 11



NNARX structure from Patil et al. (2008) (where, $u(t)z^{-1} = u(t-1)$; $y(t)z^{-1} = y(t-1)$; nk is time delay from input to output; na is the order of output polynomials; nb is the order of input polynomials; σ is constant; a, b are polynomial coefficients).

Patil et al. (2008) established the seasonal and general model for the greenhouse air temperature under tropical conditions of Thailand by NNARX. The performance of the model was evaluated by establishing regression equation between the real and the predicted data. Seen from Table 9, compared with the ARX and ARMAX models, it is found that the predicted results of NNARX are basically consistent with the real data. Frausto and Pieters (2004) also carried out a similar study, and the simulation results indicated that the model has excellent long-term performance without frequent parameter adjustment. It was found that the number of neurons in the hidden layers of the NNARX system played an essential role in achieving excellent performance.

alt-text: Table 9

Table 9

The table layout displayed in this section is not how it will appear in the final version. The representation below is solely purposed for providing corrections to the table. To preview the actual presentation of the table, please view the Proof.

Linear regression of global ARMAX ARX and NNARX models resulted over the original data for a year from Patil et al. (2008).

Season	ARX			ARMAX			NNARX		
	a	b	r^2	a	b	r^2	a	b	r^2
R	0.893	2.864	0.832	0.945	1.058	0.838	0.875	3.969	0.870
M-R	0.817	5.014	0.916	0.817	5.059	0.910	0.919	2.186	0.913
W	1.013	−0.714	0.947	1.013	−0.718	0.943	1.051	−0.250	0.962
M-W	0.979	0.070	0.959	0.989	−0.084	0.958	1.128	−4.659	0.912
S	0.972	1.264	0.952	0.976	1.131	0.951	1.063	−2.143	0.973
M-S	1.027	−0.855	0.976	1.029	−0.918	0.973	1.016	−0.260	0.968

(Notes: a is direction coefficients; r^2 is coefficients of determination; R is rainy (1–7 June); M-R is mid-rainy (1–7 August); W is winter (1–8 November); M-W is mid-winter (1–7 February); S is summer (1–7 March); M-S is midsummer (1–7 April)).

Although ANN model has a strong ability to fit nonlinear characteristics of the greenhouse, the convergence speed is slow due to the random selection of initial parameter. Fuzzy logic could solve this issue through the ability of handling both numerical data and linguistic information. Yousefi et al. (2010) designed a Neuro-Fuzzy model to predict microclimate of the greenhouse. For the model, the rule base of fuzzy approach was generated by nearest neighborhood method, and the training process of neural network was optimized by the fuzzy cluster centers. In addition, Balmat et al. (2019) and Hernández-Salazar et al. (2019) used Adaptive-Network-Based Fuzzy Inference System (ANFIS) to evaluate the evapotranspiration of greenhouse crop. ANFIS inherits interpretability of fuzzy inference system and learning ability of adaptive network. The system parameters could be changed according to the prior knowledge to make the system output closer to the real output. The result presented that data requirement was reduced by this model under the precondition of maintaining good accuracy.

In addition, Barak and Sadegh (2016) proposed three patterns to integrate the Autoregressive Integrated Moving-Average (ARIMA) and ANFIS model for forecasting energy consumption. For the first pattern, the ARIMA model was used to predict the rows of samples; then, the nonlinear capability of ANFIS was used to predict the residuals generated by the results of ARIMA model. Finally, the final model output was obtained by adding the results of ARIMA and ANFIS. For the second pattern, the forecasting of ARIMA was directly taken as one of the input features to ANFIS model, which improved the performance of ANFIS model. In order to save data, the third pattern applied

AdaBoost method with six different ANFIS structures to increase data variation. The result of MSE of AdaBoost method was decreased to 0.026% from 0.058% of the second pattern.

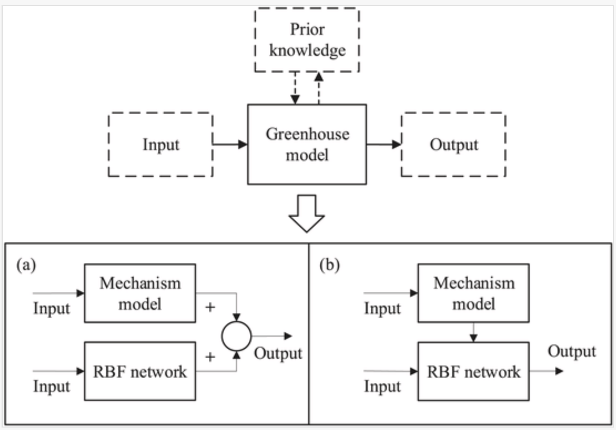
In order to reduce destructive measurements for training crop dry weight model, Soundiran et al. (2019) introduced the bootstrap resampling method into constructing process of model. The final model was obtained by weighting the outputs of multiple neural network training with different resampling datasets. The bootstrap method improved model performance under the small sample size. In addition, crop model was coupled with a greenhouse climate model, which was built by pruned ANN model, and Optimal Brain Surgeon (OBS) algorithm was used to optimize topology structure of the climate model. The greenhouse agro-ecosystem was composited with above models.

4.4.2 Integration with prior knowledge

For the greenhouse system, the machine learning model relies too much on data and ignores the critical role of prior knowledge. When the weather condition or control policy is changed, the model will get terrible prediction results. Namely, a single block-box model has a lousy extrapolation property. In order to better use prior knowledge, Linker and Seginer (2004) stated two solutions: The first strategy directly integrated the GMM and GBM; for the second strategy, the GBM trained with site data and synthetic data generated with a GMM. The hybrid physical-RBF model and prior-K sigmoid model were established, respectively, for the greenhouse temperature based on the above criteria (see Fig. 12). Because the physical model only provided prior knowledge in hybrid model and did not need to be very precise, the static model was used for modeling greenhouse inside temperature. After comparison, the hybrid and prior-K sigmoid models were not as accurate as single sigmoid neural network in the same training and verification period, but they demonstrated excellent extrapolation properties. Actually, 20 years ago, Linker et al. (1998) had modeled greenhouse temperature and humidity by physical-RBF hybrid method, and combined with the robust Failure Detection and Identification (FDI) theory of nonlinear system. The results illustrated the ability of the method to correctly detect and identify, under most circumstances, the failures were considered. Eredics and Dobrowiecki (2013) replaced the physical sub-model of Physical-RBF method with simple Multilayer Perceptron (MLP) network and compared performances of above models. As a result, MLP-RBF model was more accurate in the case that input ranges does not exceed the training domain. When the data that the model has not memorized were predicted, Physical-RBF model had an excellent performance.

alt-text: Fig. 12

Fig. 12



Methods for integrating prior knowledge into the greenhouse model to describe the work of Linker and Seginer (2004): (a) physical-RBF model; (b) prior-K sigmoid model.

In recent years, prior knowledge which was related to greenhouse system or crop has integrated into black-box model further. In some literature, mechanism/physical and black-box models were also called knowledge-driven and data-driven model, respectively. Therefore, Fan et al. (2015) proposed a plant growth model based on “knowledge” and "data". The knowledge-driven sub-model was constructed with GreenLab, which is a generic and mechanistic functional–structural plant model. GreenLab could quickly calculate the biomass accumulation of organs at each growth cycle. Meanwhile, the data-driven model was presented with RBFNN to predict dry weight of plant. Researchers mentioned that coupling methods of "knowledge" and "data" are multitudinous and particular. Prior knowledge could be integrated data-driven model by constraint function, grammar, rules, and even physically-based models. In this research, "superposition" and "composition" were used to integrate GreenLab and RBFNN model, respectively. The integrated model obtained a superior predictive ability for tomato dataset.

With the complexity of the greenhouse system, Eredics and Dobrowiecki (2010) claimed that the whole system should be divided into several sub-systems of holding respective functions. Thus, different sub-systems could be modeled by different methods. Fan et al. (2018) divided closed ecological life support system into plant and crew cabin. Where, the plant sub-system was simulated by knowledge-driven model (GreenLab and TomSim) to describe plant photosynthesis, respiration, and assimilation partitioning. The crew sub-system was built by data-driven model (Piecewise Linear Model (PLM)) to present dynamics of CO_2 production and O_2 consumption. Furthermore, above two sub-models were coupled into the mass balance model for CO_2/O_2 concentrations with metabolic stoichiometries. The result demonstrated that the coupled model has an accurate performance for predicting dry weight of plant and explain the process of material flows in a certain extent. Kim et al. (2017) pointed out that knowledge-driven and data-driven models have some limitations, respectively. The data-driven model cannot deal with unexpected and changed circumstances. The knowledge-driven model depends on the degree of understanding for system, and works under idealized assumptions and constraints. Hence, the complementary cooperation between "data" and "knowledge" is necessary. Researchers used physical laws to simulate the plant model and employed ANN to build the controller model. Then, the optimal controller model for temperature inside greenhouse was constructed by above models, and the performance was enhanced compared with existing method.

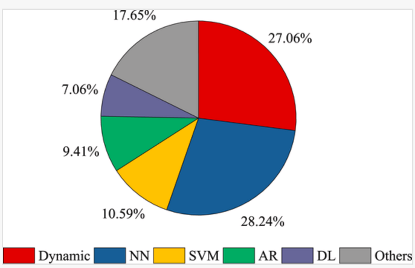
5 Discussion

As the above results, up to 73 articles that focused on modeling and solving optimization problems related to uncertainty, precision, and generalization have been reviewed to assess the main trends in the agricultural greenhouse. This literature disclosed how to establish the model for a more complex greenhouse system. Obviously, this has made design, control, and management of agricultural greenhouse better. Thereby the resource-use efficiency (e.g. water, soil, and energy) and crop productivity have been increasing, which is conducive to the development of controlled environment agriculture and agroecology. This discussed literature ware summarized in Table A.1. Among them, 73% of the literature seemed temperature and humidity as the research targets. The reason is temperature and humidity are the most critical parameters in the greenhouse climate, directly affecting the yield of crops. At the same time, they are also the primary energy consumption variables in a series of environmental factors in the greenhouse. Since the time delay of the greenhouse system, the sampling time interval of 54% summarized literature is counted by the minute, which not only avoids the expensive computational cost of a large amount of data but also reflects the changing trend of the greenhouse system.

About modeling (see Fig. 13), dynamic and neural network are favored by researchers as mainstream modeling methods, and 27% and 28% of the literature used these techniques, respectively. About optimization techniques, the heuristics optimization algorithm is widely used in the parameter identification of GMM and the optimization of GBM according to its excellent global search ability and convergence speed. Among the discussed literature, 22% and 15% of articles used PSO and EA strategy, respectively. From the latest studies from 2018 to 2020, deep learning techniques (i.e., RNN/LSTM) are increasingly applicated in agriculture greenhouse modeling, and the ability to fit complex systems and to mine spatio-temporal relationships has been proved to be better than traditional machine learning methods. In the meantime, the application of heuristics algorithms in model optimization has never diminished.

alt-text: Fig. 13

Fig. 13



Frequency for modeling methods (NN is neural network; SVM is support vector machine; AR is autoregression; DL is deep learning).

Specifically, there is a large number of uncertain parameters in the GMM, according to the complexity of the greenhouse system. For better determining these parameters, the system parameter identification has been used. Different perspectives for solving and improving parameter identification were generated based on the different components of the state-space model abstracted by the dynamic model: 1) According to a fact that there is no natural coupling relationship between the greenhouse system and actuator component, the uncertain parameters was taken apart to identify step by step, which reduces the complexity of identification; 2) Multi-objective optimization was applied in parameter identification to solve the multi-output characteristics of greenhouse system; 3) In order to reduce the identification speed in high dimensional parameter space, sensitivity-analysis was pointed out to reduce the dimension of uncertain parameters; 4) The online identification method was proposed to solve the time-varying and hysteresis of greenhouse system. Besides, GBM has been usually used to model the complex and nonlinear greenhouse system, which effectively improved the accuracy of the model. To eliminate the randomness and blindness of model hyperparameter selection and improve the model training speed, heuristic algorithms have been proposed to optimize GBM. According to the reviewed literature, the optimization of the model mainly focused on the adjustment of SVM hyperparameter, BPNN training process, and RBFNN model structure and parameters.

Although GBM has excellent accuracy for complex systems, it is too data-dependent to be easily adjusted. GMM has good explanatory ability and flexibility, but it is not easy to model complex systems. Since the application of different models has its own advantages and disadvantages, it is difficult to determine which modeling method is the best. The idea of multi-model integration was proposed in discussed literature, which integrated different architectures GBM, and the accuracy of the model can be further improved. Through the method of fusing the GMM and GBM into the ensemble model, namely prior knowledge was integrated into the GBM, which could effectively enhance the extrapolation properties of the model.

6 Conclusion

An excellent greenhouse model (e.g. microclimate or crop behavior model) could help environmental control more advanced, systematic management more timely and precise, and structural design more convenient. In this situation, resource-use efficiency is increased, external inputs of fertilizer and pesticides are reduced, and cleaner crop productivity and sustainability are promoted. Therefore, this paper intensively reviews the existing literature concerning modeling and improvement methods in the agricultural greenhouse system. In the meanwhile, interested approaches, design specifications, and technical trends in the field of greenhouse modeling were extracted. The main valuable and concise conclusions are as follows:

- Most literature consider the temperature and humidity as the main target of modeling, which directly affect crop production and assist in the design of low energy consumption control strategies.
- The types of existing techniques for modeling greenhouse were classified as mechanism method, time series method, and machine learning method. This classification approach is based on the structure and principle on which the model depends. In particular, dynamic and neural network were more adopted, and the application of deep learning to time series prediction tasks of the greenhouse was increased in recent years.
- According to problem-orientation, the methods of improvement and optimization of the model are summarized as 1) system parameter identification for the dilemma of parameter uncertainty of mechanism model, 2) structure and process optimization for the dilemma of random hyperparameters selection, slow convergence, and local optimum of the model, and 3) multi-model integration for the dilemma of the single model with awful generalization and extrapolation properties as well as low accuracy.
- For most studies on greenhouse modeling, the role of crops has not been highlighted, and the assumptions of the model are relatively harsh. 53% of the summarized literature has taken crops as a component in greenhouse system and 24% considered the effect of crops in modeling. However, in the modern application of large-scale production and vertical greenhouse, the transpiration and photosynthesis of crops will have a significant impact on the greenhouse system. Besides, the particular greenhouse system, such as the greenhouse aquaponics system, cannot be well promoted.

In the future, with the development of big data and artificial intelligence, more advanced machine learning technology (e.g. representation learning, transfer learning, online learning, etc.) will be widely used in engineering. Especially, the importance of stable learning by causal inference and fusion with prior knowledge should be highlighted to achieve better extrapolation properties. Furthermore, there are few multi-model studies on the object of agricultural greenhouse directly. It is a valuable research topic on how to integrate multiple different models better and maximize the application of existing prior information to increase the “transparency” of greenhouse black-box model, namely fusing “knowledge” and “data".

Our project of Environment Optimization and Regulation of Multi-temporal-spatial-scale Greenhouses from Overseas High-level Youth Talents Program has been investigated by the present authors. Optimization and regulation of greenhouse environment for a novel and cleaner food production system – aquaponics in greenhouses will be carried out soon.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was financially supported by Overseas High-level Youth Talents Program (China Agricultural University, China, Grant No. 62339001), Science and Technology Cooperation – Sino-Malta Fund 2019: Research and Demonstration of Real-time Accurate Monitoring System for Early-stage Fish in Recirculating Aquaculture System (AquaDetector, Grant No. ~~SQ~~2019YFE010~~1423~~3700), China Agricultural University Excellent Talents Plan (Grant No. 31051015), Major Science and Technology Innovation Fund 2019 of Shandong Province (Grant No. 2019JZZY010703), National Innovation Center for Digital Fishery, and Beijing Engineering and Technology Research Center for Internet of Things in Agriculture. The authors also appreciate constructive and valuable comments provided by reviewers.

Appendix A Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jclepro.2020.124843.

Nomenclature

Abbreviations

ANFIS Adaptive Neuro-Fuzzy Interference System
ANN Artificial Neural Network
APSO Adaptive Particle Swarm Optimization
ARIMA Autoregressive Integrated Moving-Average
ARMAX Autoregressive Moving-Average with Exogenous
ARX Autoregressive with Exogenous
BP Back Propagation
BPNN Back Propagation Neural Network
CFD Computational Fluid Dynamics
DE Differential Evolution
ELM Extreme Learning Machine
EA Evolutionary Algorithm
EP Evolution Programming
ES Evolution Strategy
FFNN Feed Forward backpropagation Neural Network
GA Genetic Algorithm
GBM Greenhouse Black-box Model
GMLM Greenhouse Machine Learning Model
GMM Greenhouse Mechanism Model
GTSM Greenhouse Time Series Model
IBP Improved Back Propagation
ICFPSO Increased Convergence Factor Particle Swarm Optimization
IPSO Improved Particle Swarm Optimization
KELM Kernel based Extreme Learning Machine
KNN k-Nearest Neighbor
LM Levenberg-Marquardt
LSQ Least Squares
LSSVM Least-Squares Support Vector Machine
LSTM Long Short-Term Memory
MLP Multilayer Perceptron
MLR Multiple Linear Regression
MOEA Multi-Objective Evolutionary Algorithm
MOGA Multi-Objective Genetic Algorithm
MOLPSO Many Optimizing Liaisons Particle Swarm Optimization
NARX Nonlinear Autoregressive Exogenous
NNARX Neural Network Autoregressive with Exogenous

OBF Orthonormal Basis Function
OBS Optimal Brain Surgeon
OLS Orthogonal Least Squares
OS Online Sparse
PCA Principal Component Analysis
PLM Piecewise Linear Mode
PO Process Optimization
POD Proper Orthogonal Decomposition
PSO Particle Swarm Optimization
RBFNN Radial Basis Function Neural Network
RLSEF Recursive Least Squares Algorithm with Exponential Forgetting
RNN Recurrent Neural Network
ROLS Regularized Orthogonal Least Squares
SA Sensitivity Analysis
SO Structural Optimization
SPI System Parameter Identification
SQP Sequential Quadratic Programming
SVM Support Vector Machine
UPs Uncertain-Parameters

Symbols

A_{co_2} CO2 injecting
A_{ven} Natural ventilation/window opening
A_{fan} Mechanical ventilation/The activation of fans
A_{heat} Heating air and soil
A_{fog} Fogging
A_{cool} Cooling
A_{sha} Shading/Thermal screen/Sliding shutter
A_{ir} Irrigation
DO Dissolved oxygen
E Power consumption/Energy demand/energy consumption/Heating consumption
EC EC of substrate/EC of nutrient solutions
G Growth stag/Plant height/Dry matter of crop and fruit/Number of nodes
H_R Relative humidity/Vapor pressure
H_a Absolute humidity
H_s Soil humidity/Moisture content of substrate
I Solar radiation/illumination
LAI Leaf area index
MT Mark of time (e.g. the hour of day, the day of the week)
OD Optical density
pH Potential of hydrogen
p_{cl} Cloudiness of sky
PAR Photosynthetic Active Radiation
Q Solar heat load/Heat flux distributions
R_{res} Respiration rate
R_{tran} Transpiration/evapotranspiration rate
R_{pho} Photosynthesis rate/Chlorophyll-a
T_a Air temperature
T_w Wall temperature
T_s Soil/assimilation box device/substrate temperature
T_c Cover/roof temperature
T_p Plant/crop temperature
T_{wat} Supply/return/pipe water temperature
w_v Wind speed/wind velocity
w_d Wind direction
ρ_{CO_2} CO2 concentration
ρ_{O_2} O2 concentration
ρ_{nut} Nutrient concentration (include: nitrate, phosphate, salinity, etc.)

ρ_{air} Air density

Notation

a A scalar (integer or real)

$\boldsymbol{a}$ A vector

$\dot{\boldsymbol{x}}$ Derivative of $\boldsymbol{x}$ with respect to t

$\|\boldsymbol{x}\|_2$ L^2 norm of $\boldsymbol{x}$

$f(\boldsymbol{x};\boldsymbol{\theta})$ A function of $\boldsymbol{x}$ parametrized by $\boldsymbol{\theta}$

$J(\boldsymbol{\theta};k)$ A cost function of k-th sample parametrized by $\boldsymbol{\theta}$

$\hat{\boldsymbol{y}}$ An output vector of model

Uncited references

~~Yin et al. (2018)~~.

References

The corrections made in this section will be reviewed and approved by a journal production editor. The newly added/removed references and its citations will be reordered and rearranged by the production team.

Ahamed, M.S., Guo, H., Tanino, K., 2019. Energy saving techniques for reducing the heating cost of conventional greenhouses. Biosyst. Eng. 178, 9–33. doi:10.1016/j.biosystemseng.2018.10.017.

Aiello, G., Giovino, I., Vallone, M., Catania, P., Argento, A., 2018. A decision support system based on multisensor data fusion for sustainable greenhouse management. J. Clean. Prod. 172, 4057–4065. doi:10.1016/j.jclepro.2017.02.197.

Ali, A., Hassanein, H.S., 2019. Wireless sensor network and deep learning for prediction greenhouse environments. In: 2019 International Conference on Smart Applications, Communications and Networking (SmartNets). Presented at the 2019 International Conference on Smart Applications, Communications and Networking. SmartNets), pp. 1–5. doi:10.1109/SmartNets48225.2019.9069766.

Ángel, G.-N.M., Raquel, M.-E., Andrés, B.-C., Belén, A., Luis, M.J., M, C.J., 2019. An LSTM deep learning scheme for prediction of low temperatures in agriculture. Ambient Intell. Smart Environ. 130–138. doi:10.3233/AISE190032.

Baddadi, S., Bouadila, S., Ghorbel, W., Guizani, A., 2019. Autonomous greenhouse microclimate through hydroponic design and refurbished thermal energy by phase change material. J. Clean. Prod. 211, 360–379. doi:10.1016/j.jclepro.2018.11.192.

Balmat, J.-F., Lafont, F., Ali, A.M., Pessel, N., Fernández, J.C.R., 2019. Evaluation of the reference evapotranspiration for a greenhouse crop using an Adaptive-Network-Based Fuzzy Inference System (ANFIS). In: Proceedings of the 3rd International Conference on Machine Learning and Soft Computing, ICMLSC 2019. Association for Computing Machinery, Da Lat, Viet Nam, pp. 203–205. doi:10.1145/3310986.3310987.

Bao, Y., Hu, Z., Xiong, T., 2013. A PSO and pattern search based memetic algorithm for SVMs parameters optimization. Neurocomputing 117, 98–106. doi:10.1016/j.neucom.2013.01.027.

Barak, S., Sadegh, S.S., 2016. Forecasting energy consumption using ensemble ARIMA–ANFIS hybrid algorithm. Int. J. Electr. Power Energy Syst. 82, 92–104. doi:10.1016/j.ijepes.2016.03.012.

Blasco, X., Martínez, M., Herrero, J.M., Ramos, C., Sanchis, J., 2007. Model-based predictive control of greenhouse climate for reducing energy and water consumption. Comput. Electron. Agric. 55, 49–70. doi:10.1016/j.compag.2006.12.001.

Businger, J.A., 1963. The Glasshouse Climate, Physics of Plant Environment. Van Wijk North Publ Co Amst, pp. 270–318.

Chapelle, O., Vapnik, V., Bousquet, O., Mukherjee, S., 2002. Choosing multiple parameters for support vector machines. Mach. Learn. 46, 131–159. doi:10.1023/A:1012450327387.

Chen, J., Yang, J., Zhao, J., Xu, F., Shen, Z., Zhang, L., 2016. Energy demand forecasting of the greenhouses using nonlinear models based on model optimized prediction method. Neurocomputing 174, 1087–1100. doi:10.1016/j.neucom.2015.09.105.

Chen, S., Chng, E.S., Alkadhimi, K., 1996. Regularized orthogonal least squares algorithm for constructing radial basis function networks. Int. J. Contr. 64, 829–837. doi:10.1080/00207179608921659.

Choab, N., Allouhi, A., El Maakoul, A., Kousksou, T., Saadeddine, S., Jamil, A., 2019. Review on greenhouse microclimate and application: design parameters, thermal modeling and simulation, climate controlling technologies. Sol. Energy 191, 109–137. doi:10.1016/j.solener.2019.08.042.

Coelho, J.P., de Moura Oliveira, P.B., Cunha, J.B., 2005. Greenhouse air temperature predictive control using the particle swarm optimisation algorithm. Comput. Electron. Agric. 49, 330–344 Modelling and control in agricultural processes. doi:10.1016/j.compag.2005.08.003.

Cruz-Valeriano, E., Begovich, O., Ruiz-Leon, J., 2013. Modeling of a greenhouse using particle swarm optimization. In: 2013 10th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE). Presented at the 2013 10th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE). IEEE, Mexico City, Mexico, pp. 268–273. doi:10.1109/ICEEE.2013.6676034.

Cunha, J.B., 2003. Greenhouse climate models: an overview. In: Efita 2003 Conference. pp. 823–829.

Elanchezhian, A., Basak, J.K., Park, J., Khan, F., Okyere, F.G., Lee, Y., Bhujel, A., Lee, D., Sihalath, T., Kim, H.T., 2020. Evaluating different models used for predicting the indoor microclimatic parameters of a greenhouse. Appl. Ecol. Environ. Res. 18, 2141–2161. doi:10.15666/aeer/1802_21412161.

Eredics, P., Dobrowiecki, T.P., 2013. Hybrid MLP-RBF model structure for short-term internal temperature prediction in greenhouse environments. In: 2013 IEEE 14th International Symposium on Computational Intelligence and Informatics (CINTI). Presented at the 2013 IEEE 14th International Symposium on Computational Intelligence and Informatics (CINTI). pp. 377–380. doi:10.1109/CINTI.2013.6705225.

Eredics, P., Dobrowiecki, T.P., 2010. Hybrid knowledge modeling for an intelligent greenhouse. In: IEEE 8th International Symposium on Intelligent Systems and Informatics. Presented at the IEEE 8th International Symposium on Intelligent Systems and Informatics. pp. 459–463. doi:10.1109/SISY.2010.5647334.

Escamilla-García, A., Soto-Zarazúa, G.M., Toledano-Ayala, M., Rivas-Araiza, E., Gastélum-Barrios, A., 2020. Applications of artificial neural networks in greenhouse technology and overview for smart agriculture development. Appl. Sci. Switz. 10. doi:10.3390/app10113835.

Esen, H., Esen, M., Yuksel, T., 2015. Modelling of biogas, solar and a ground source heat pump greenhouse heating system by using ensemble learning. New Dev. Mech. Mech. Eng. 74–81.

Esmaeli, H., Roshandel, R., 2020. Optimal design for solar greenhouses based on climate conditions. Renew. Energy 145, 1255–1265. doi:10.1016/j.renene.2019.06.090.

Fan, X.-R., Kang, M.-Z., Heuvelink, E., de Reffye, P., Hu, B.-G., 2015. A knowledge-and-data-driven modeling approach for simulating plant growth: a case study on tomato growth. Ecol. Model. 312, 363–373. doi:10.1016/j.ecolmodel.2015.06.006.

Fan, X.-R., Wang, X., Kang, M., Hua, J., Guo, S., de Reffye, P., Hu, B.-G., 2018. A knowledge-and-data-driven modeling approach for simulating plant growth and the dynamics of CO2/O2 concentrations in a closed system of plants and humans by integrating mechanistic and empirical models. Comput. Electron. Agric. 148, 280–290. doi:10.1016/j.compag.2018.03.006.

FAO, 2018. The 10 Elements of Agroecology: Guiding the Transition to Sustainable Food and Agricultural Systems. FAO, Rome, Italy.

Ferreira, P.M., Ruano, A.E., 2008. Application of computational intelligence methods to greenhouse environmental modelling. In: 2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence). Presented at the 2008 IEEE International Joint Conference on Neural Networks (IJCNN 2008 - Hong Kong). IEEE, Hong Kong, China, pp. 3582–3589. doi:10.1109/IJCNN.2008.4634310.

Fox, J.A., Adriaanse, P., Stacey, N.T., 2019. Greenhouse energy management: the thermal interaction of greenhouses with the ground. J. Clean. Prod. 235, 288–296. doi:10.1016/j.jclepro.2019.06.344.

Frausto, H.U., Pieters, J.G., 2004. Modelling greenhouse temperature using system identification by means of neural networks. Neurocomputing 56, 423–428. doi:10.1016/j.neucom.2003.08.001.

García Nieto, P.J., García-Gonzalo, E., Alonso Fernández, J.R., Díaz Muñiz, C., 2016. A hybrid PSO optimized SVM-based model for predicting a successful growth cycle of the Spirulina platensis from raceway experiments data. J. Comput. Appl. Math. 291, 293–303 Mathematical Modeling and Computational Methods. doi:10.1016/j.cam.2015.01.009.

Grabarczyk, S., 2018. Modeling of heat consumption in a greenhouse using experimental data. In: Presented at the E3S Web of Conferences. doi:10.1051/e3sconf/20184900037.

Guo, L.-L., Li, Y., Liu, Z.-W., Sun, H.-W., 2019. Application of particle swarm optimization BP algorithm in air humidity of greenhouse crops. In: IOP Conference Series: Materials Science and Engineering. IOP Publishing, Sanya, China, 012003. doi:10.1088/1757-899X/565/1/012003.

Guzmán-Cruz, R., Castañeda-Miranda, R., García-Escalante, J.J., López-Cruz, I.L., Lara-Herrera, A., de la Rosa, J.I., 2009. Calibration of a greenhouse climate model using evolutionary algorithms. Biosyst. Eng. 104, 135–142. doi:10.1016/j.biosystemseng.2009.06.006.

Guzmán-Cruz, R., Olvera-González, E., López-Cruz, I.L., Montoya-Zamora, R., 2013. Modelling greenhouse air temperature using evolutionary algorithms in auto regressive models. Afr. J. Agric. Res. 8, 251–261. doi:10.5897/AJAR12.280.

Hasni, A., Draoui, B., Latfaoui, M., Boulard, T., 2010. Identification of natural ventilation parameters in a greenhouse with continuous roof vents, using a PSO and GAs. Sens. Transducers 119, 182.

Hasni, A., Taibi, R., Draoui, B., Boulard, T., 2011. Optimization of greenhouse climate model parameters using particle swarm optimization and genetic algorithms. Energy Procedia 6, 371–380. doi:10.1016/j.egypro.2011.05.043.

He, F., Ma, C., 2010. Modeling greenhouse air humidity by means of artificial neural network and principal component analysis. In: Comput. Electron. Agric., Special Issue on Computer and Computing Technologies in Agriculture 71. pp. S19–S23. doi:10.1016/j.compag.2009.07.011.

He, F., Ma, C., Zhang, J., Chen, Y., 2007. Greenhouse air temperature and humidity prediction based on improved BP neural network and genetic algorithm. In: Liu, D., Fei, S., Hou, Z., Zhang, H., Sun, C. (Eds.), Advances in Neural Networks – ISNN 2007, Lecture Notes in Computer Science. Springer, Berlin, Heidelberg, pp. 973–980. doi:10.1007/978-3-540-72395-0_119.

Hernández-Salazar, J.A., Hernández-Rodríguez, D., Hernández-Cruz, R.A., Ramos-Fernández, J.C., Márquez-Vera, M.A., Trejo-Macotela, F.R., 2019. Estimation of the evapotranspiration using ANFIS algorithm for agricultural production in greenhouse. In: 2019 IEEE International Conference on Applied Science and Advanced Technology (ICASAT). Presented at the 2019 IEEE International Conference on Applied Science and Advanced Technology (iCASAT). pp. 1–5. doi:10.1109/iCASAT48251.2019.9069533.

Herrero, J.M., Blasco, X., Martínez, M., Ramos, C., Sanchis, J., 2008. Robust identification of non-linear greenhouse model using evolutionary algorithms. Contr. Eng. Pract. 16, 515–530. doi:10.1016/j.conengprac.2007.06.001.

Herrero, J.M., Blasco, X., Martínez, M., Ramos, C., Sanchis, J., 2007. Non-linear robust identification of a greenhouse model using multi-objective evolutionary algorithms. Biosyst. Eng. 98, 335–346. doi:10.1016/j.biosystemseng.2007.06.004.

Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. Neural Comput. 9, 1735–1780. doi:10.1162/neco.1997.9.8.1735.

Huang, C.-L., Dun, J.-F., 2008. A distributed PSO–SVM hybrid system with feature selection and parameter optimization. Appl. Soft Comput. 8, 1381–1391. doi:10.1016/j.asoc.2007.10.007.

Huang, C.-L., Wang, C.-J., 2006. A GA-based feature selection and parameters optimization for support vector machines. Expert Syst. Appl. 31, 231–240. doi:10.1016/j.eswa.2005.09.024.

Hui, Z., Lin-lin, Q., Gang, W., 2017. Modeling and simulation of greenhouse temperature hybrid system based on ARMAX model. In: 2017 36th Chinese Control Conference (CCC). Presented at the 2017 36th Chinese Control Conference (CCC). pp. 2237–2241. doi:10.23919/ChiCC.2017.8027690.

Iddio, E., Wang, L., Thomas, Y., McMorrow, G., Denzer, A., 2020. Energy efficient operation and modeling for greenhouses: a literature review. Renew. Sustain. Energy Rev. 117, 109480. doi:10.1016/j.rser.2019.109480.

Joudi, K.A., Farhan, A.A., 2015. A dynamic model and an experimental study for the internal air and soil temperatures in an innovative greenhouse. Energy Convers. Manag. 91, 76–82. doi:10.1016/j.enconman.2014.11.052.

Jovanović, R.Ž., Sretenović, A.A., Živković, B.D., 2015. Ensemble of various neural networks for prediction of heating energy consumption. Energy Build. 94, 189–199. doi:10.1016/j.enbuild.2015.02.052.

Jung, D.-H., Kim, H.S., Jhin, C., Kim, H.-J., Park, S.H., 2020. Time-serial analysis of deep neural network models for prediction of climatic conditions inside a greenhouse. Comput. Electron. Agric. 173, 105402. doi:10.1016/j.compag.2020.105402.

Kamal, S., 2013. The Renewable Revolution: How We Can Fight Climate Change, Prevent Energy Wars, Revitalize the Economy and Transition to a Sustainable Future. Routledge.

Khanlari, A., Sözen, A., Şirin, C., Tuncer, A.D., Gungor, A., 2020. Performance enhancement of a greenhouse dryer: analysis of a cost-effective alternative solar air heater. J. Clean. Prod. 251, 119672. doi:10.1016/j.jclepro.2019.119672.

Kim, B.S., Kang, B.G., Choi, S.H., Kim, T.G., 2017. Data modeling versus simulation modeling in the big data era: case study of a greenhouse control system. Simulat.Trans. Soc. Model. Simulat. Int. 93, 579–594. doi:10.1177/0037549717692866.

Kim, K., Kim, D.-K., Noh, J., Kim, M., 2018. Stable forecasting of environmental time series via long short term memory recurrent neural network. IEEE Access 6, 75216–75228. doi:10.1109/ACCESS.2018.2884827.

Lagerberg, C., Brown, M.T., 1999. Improving agricultural sustainability: the case of Swedish greenhouse tomatoes. J. Clean. Prod. 7, 421–434. doi:10.1016/S0959-6526(99)00230-9.

Laktionov, I.S., Vovna, O.V., Kabanets, M.M., Derzhevetska, M.A., Zori, A.A., 2020. Mathematical model of measuring monitoring and temperature control of growing vegetables in greenhouses. Int. J. Des. Nat. Ecodyn. 15, 325–336. doi:10.18280/ijdne.150306.

Lammari, K., Bounaama, F., Draoui, B., Mrah, B., Haidas, M., 2012. In: GA Optimization of the Coupled Climate Model of an Order Two of a Greenhouse, vol. 18. Energy Procedia, Terragreen, pp. 416–425 2012: Clean Energy Solutions for Sustainable Environment (CESSE). doi:10.1016/j.egypro.2012.05.053.

Li, F., Li, S., Wang, Z., Chen, Z., Zhao, X., 2016. Design and research of intelligent greenhouse monitoring system based on Internet of Things. In: Presented at the 2016 International Conference on Computer Science and Electronic Technology. Atlantis Press, pp. 76–79. doi:10.2991/cset-16.2016.19.

Li, J., Li, L., Wang, H., Ferentinos, K.P., Li, M., Sigrimis, N., 2017. Proactive energy management of solar greenhouses with risk assessment to enhance smart specialisation in China. Biosyst. Eng. 158, 10–22. doi:10.1016/j.biosystemseng.2017.03.007.

Li, K., Sha, Z., Xue, W., Chen, X., Mao, H., Tan, G., 2020. A fast modeling and optimization scheme for greenhouse environmental system using proper orthogonal decomposition and multi-objective genetic algorithm. Comput. Electron. Agric. 168, 105096. doi:10.1016/j.compag.2019.105096.

Li, T., Ji, Y., Zhang, M., Li, M., 2017. Determining optimal CO2 concentration of greenhouse tomato based on PSO-SVM. Appl. Eng. Agric. 33, 157–166. doi:10.13031/aea.11578.

Li, T., Ji, Y., Zhang, M., Sha, S., Li, M., 2017. Universality of an improved photosynthesis prediction model based on PSO-SVM at all growth stages of tomato. Int. J. Agric. Biol. Eng. 10, 63–73. doi:10.25165/ijabe.v10i2.2580.

Li, Z., Wang, S., Gao, C., Zhang, L., Chu, Q., 2002. Research and application of auto-control system for solar greenhouse comprehensive environment with temperature as principal parameter. Trans. Chin. Soc. Agric. Eng. 18, 68–71. doi:10.3321/j.issn:1002-6819.2002.03.017.

Liang, M.-H., Chen, L.-J., He, Y.-F., Du, S.-F., 2018. Greenhouse temperature predictive control for energy saving using switch actuators. In: IFAC-pap., 6th IFAC Conference on Bio-Robotics BIOROBOTICS 2018, vol. 51. pp. 747–751. doi:10.1016/j.ifacol.2018.08.106.

Lin, D., Wei, R., Xu, L., 2019. An integrated yield prediction model for greenhouse tomato. Agronomy 9, 873. doi:10.3390/agronomy9120873.

Lin, D., Zhang, L., Xia, X., 2020. Hierarchical model predictive control of Venlo-type greenhouse climate for improving energy efficiency and reducing operating cost. J. Clean. Prod. 264, 121513. doi:10.1016/j.jclepro.2020.121513.

Linker, R., Gutman, P.O., Seginer, I., 1998. Robust failure detection and identification in a greenhouse modeled with hybrid physical/neural network models. IFAC Proc. 31, 173–178. doi:10.1016/S1474-6670(17)36060-3.

Linker, R., Seginer, I., 2004. Greenhouse temperature modeling: a comparison between sigmoid neural networks and hybrid models. Math. Comput. Simulat. 65, 19–29 Selected papers of the IMACS/IFAC Fourth International Symposium on Mathematical Modelling and Simulation in Agricultural and Bio-Industries. doi:10.1016/j.matcom.2003.09.004.

Liu, Q., Jin, D., Shen, J., Fu, Z., Linge, N., 2016. A WSN-based prediction model of microclimate in a greenhouse using extreme learning approaches. In: 2016 18th International Conference on Advanced Communication Technology (ICACT). Presented at the 2016 18th International Conference on Advanced Communication Technology (ICACT). IEEE, Pyeongchang Kwangwoon Do, South Korea, pp. 730–735. doi:10.1109/ICACT.2016.7423609.

Liu, Y., Gong, C., Yang, L., Chen, Y., 2020. DSTP-RNN: a dual-stage two-phase attention-based recurrent neural network for long-term and multivariate time series prediction. Expert Syst. Appl. 143, 113082. doi:10.1016/j.eswa.2019.113082.

Liu, Y., Zhang, Q., Song, L., Chen, Y., 2019. Attention-based recurrent neural networks for accurate short-term and long-term dissolved oxygen prediction. Comput. Electron. Agric. 165, 104964. doi:10.1016/j.compag.2019.104964.

Ljung, L., 1978. Convergence analysis of parametric identification methods. IEEE Trans. Automat. Contr. 23, 770–783. doi:10.1109/TAC.1978.1101840.

Llera, J.R., Deb, K., Runkle, E., Xu, L., Goodman, E., 2019. Evolving and comparing greenhouse control strategies using model-based multi-objective optimization. In: Presented at the Proceedings of the 2018 IEEE Symposium Series on Computational Intelligence, SSCI 2018. pp. 1929–1936. doi:10.1109/SSCI.2018.8628783.

López-Cruz, I.L., Fitz-Rodríguez, E., Salazar-Moreno, R., Rojano-Aguilar, A., Kacira, M., 2018. Development and analysis of dynamical mathematical models of greenhouse climate: a review. Eur. J. Hortic. Sci. 83, 269–279. doi:10.17660/eJHS.2018/83.5.1.

McNulty, J., 2017. Solar greenhouses generate electricity and grow crops at the same time. UC St. Cruz Mag. Novemb. 03.

Mirzamohammadi, S., Jabarzadeh, A., Shahrabi, M.S., 2020. Long-term planning of supplying energy for greenhouses using renewable resources under uncertainty. J. Clean. Prod. 264, 121611. doi:10.1016/j.jclepro.2020.121611.

Mohammadi, B., Ranjbar, S.F., Ajabshirchi, Y., 2018. Application of dynamic model to predict some inside environment variables in a semi-solar greenhouse. Inf. Process. Agric. 5, 279–288. doi:10.1016/j.inpa.2018.01.001.

Moon, T., Ahn, T.I., Son, J.E., 2018. Forecasting root-zone electrical conductivity of nutrient solutions in closed-loop soilless cultures via a recurrent neural network using environmental and cultivation information. Front. Plant Sci. 9, 859. doi:10.3389/fpls.2018.00859.

Outanoute, M., Lachhab, A., Selmani, A., Oubehar, H., Snoussi, A., Guerbaoui, M., Ed-dahhak, A., Bouchikhi, B., 2018. Particle swarm optimization of BP-ANN based soft sensor for greenhouse climate. J. Electron. Commer. Organ. JECO 16, 72–81. doi:10.4018/jeco.2018010106.

Patil, S.L., Tantau, H.J., Salokhe, V.M., 2008. Modelling of tropical greenhouse temperature by auto regressive and neural network models. Biosyst. Eng. 99, 423–431 https://doi.org/10/d7tkzx.

Pérez-González, A., Begovich, O., Ruiz-Leon, J., 2014. Modeling of a greenhouse prototype using PSO algorithm based on a LabView™ application. In: 2014 11th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE). Presented at the 2014 11th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE). IEEE, Ciudad del Carmen, Campeche, Mexico, pp. 1–6. doi:10.1109/ICEEE.2014.6978281.

Pérez-González, A., Begovich-Mendoza, O., Ruiz-León, J., 2018. Modeling of a greenhouse prototype using PSO and differential evolution algorithms based on a real-time LabView™ application. Appl. Soft Comput. 62, 86–100. doi:10.1016/j.asoc.2017.10.023.

Qin, Y., Song, D., Chen, H., Cheng, W., Jiang, G., Cottrell, G.W., 2017. A dual-stage attention-based recurrent neural network for time series prediction. In: Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, vol. 17. IJCAI, pp. 2627–2633. doi:10.24963/ijcai.2017/366.

Rhinehart, R.R., 2016. Nonlinear Regression Modeling for Engineering Applications: Modeling, Model Validation, and Enabling Design of Experiments. John Wiley & Sons.

Avissar, Roni, Mahrer, Ytzhaq, 1982. Verification study of a numerical greenhouse microclimate model. Trans. ASAE (Am. Soc. Agric. Eng.) 25, 1711–1720. doi:10.13031/2013.33794.

Saberian, A., Sajadiye, S.M., 2019. The effect of dynamic solar heat load on the greenhouse microclimate using CFD simulation. Renew. Energy 138, 722–737. doi:10.1016/j.renene.2019.01.108.

Salah, L.B., Fourati, F., 2020. Deep elman neural network for greenhouse modeling. In: Bouhlel, M.S., Rovetta, S. (Eds.), Proceedings of the 8th International Conference on Sciences of Electronics, Technologies of Information and Telecommunications, Smart Innovation, Systems and Technologies. Springer, Cham, pp. 271–280. doi:10.1007/978-3-030-21005-2_26.

Salgado, P., Cunha, J.B., 2005. Greenhouse climate hierarchical fuzzy modelling. Contr. Eng. Pract. 13, 613–628. doi:10.1016/j.conengprac.2004.05.007.

Sethi, V.P., Sumathy, K., Lee, C., Pal, D.S., 2013. Thermal modeling aspects of solar greenhouse microclimate control: a review on heating technologies. Sol. Energy 96, 56–82. doi:10.1016/j.solener.2013.06.034.

Shen, Y., Wei, R., Xu, L., 2018. Energy consumption prediction of a greenhouse and optimization of daily average temperature. Energies 11, 65. doi:10.3390/en11010065.

Singh, M.C., Singh, J.P., Singh, K.G., 2018. Development of a microclimate model for prediction of temperatures inside a naturally ventilated greenhouse under cucumber crop in soilless media. Comput. Electron. Agric. 154, 227–238. doi:10.1016/j.compag.2018.08.044.

Singh, Mahesh Chand, Singh, J.P., Singh, K.G., 2018. Development of a microclimate model for prediction of temperatures inside a naturally ventilated greenhouse under cucumber crop in soilless media. Comput. Electron. Agric. 154, 227–238. doi:10.1016/j.compag.2018.08.044.

Singh, M.C., Yousuf, A., Singh, J., 2016. Greenhouse microclimate modeling under cropped conditions: a review. Res. Environ. Life Sci. 9, 1552–1557.

Soundiran, R., Radhakrishnan, T.K., Natarajan, S., 2019. Modeling of greenhouse agro-ecosystem using optimally designed bootstrapping artificial neural network. Neural Comput. Appl. 31, 7821–7836. doi:10.1007/s00521-018-3598-7.

Steinwart, I., Christmann, A., 2008. Support Vector Machines. Springer Science & Business Media.

Su, Y., Xu, L., 2017. Towards discrete time model for greenhouse climate control. Eng. Agric. Environ. Food 10, 157–170. doi:10.1016/j.eaef.2017.01.001.

Takakura, T., Jordan, K.A., Boyd, L.L., 1971. Dynamic simulation of plant growth and environment in the greenhouse. Trans. ASAE (Am. Soc. Agric. Eng.) 14 0964–0971. doi:10.13031/2013.38432.

Taki, M., Ajabshirchi, Y., Ranjbar, S.F., Rohani, A., Matloobi, M., 2016. Heat transfer and MLP neural network models to predict inside environment variables and energy lost in a semi-solar greenhouse. Energy Build. 110, 314–329 https://doi.org/10/f767bd.

Tovany, J.L., Ross-León, R., Ruiz-León, J., Ramírez-Treviño, A., Begovich, O., 2013. Greenhouse modeling using continuous timed petri Nets. Math. Probl Eng. 639306 2013. doi:10.1155/2013/639306.

Trabelsi, A., Lafont, F., Kamoun, M., Enea, G., 2007. Fuzzy identification of a greenhouse. Appl. Soft Comput. 7, 1092–1101. doi:10.1016/j.asoc.2006.06.009.

Trejo-Perea, M., Herrera-Ruiz, G., Rios-Moreno, J., Miranda, R.C., Rivas, E., 2009. Greenhouse energy consumption prediction using neural networks models. Int. J. Agric. Biol. 11, 6.

Uchida Frausto, H., Pieters, J.G., Deltour, J.M., 2003. Modelling greenhouse temperature by means of auto regressive models. Biosyst. Eng. 84, 147–157 https://doi.org/10/bft2df.

United Nations, Department of Economic and Social Affairs, Population Division, 2019. World Population Prospects Highlights, 2019 Revision Highlights, 2019 Revision.

Vanthoor, B.H.E., de Visser, P.H.B., Stanghellini, C., van Henten, E.J., 2011. A methodology for model-based greenhouse design: Part 2, description and validation of a tomato yield model. Biosyst. Eng. 110, 378–395 https://doi.org/10/fgzdvg.

Vanthoor, B.H.E., Stanghellini, C., van Henten, E.J., de Visser, P.H.B., 2011. A methodology for model-based greenhouse design: Part 1, a greenhouse climate model for a broad range of designs and climates. Biosyst. Eng. 110, 363–377 https://doi.org/10/ck5xjt.

Wang, D., Wang, M., Qiao, X., 2009. Support vector machines regression and modeling of greenhouse environment. Comput. Electron. Agric. 66, 46–52. doi:10.1016/j.compag.2008.12.004.

Wang, Hongkang, Li, L., Wu, Y., Meng, F., Wang, Haihua, Sigrimis, N.A., 2018. Recurrent neural network model for prediction of microclimate in solar greenhouse. In: IFAC-pap., 6th IFAC Conference on Bio-Robotics BIOROBOTICS 2018, vol. 51. pp. 790–795. doi:10.1016/j.ifacol.2018.08.099.

Wang, L., Wang, B., 2020. Construction of greenhouse environment temperature adaptive model based on parameter identification. Comput. Electron. Agric. 174. doi:10.1016/j.compag.2020.105477.

Wang, S., Boulard, T., 2000. Predicting the microclimate in a naturally ventilated plastic house in a mediterranean climate. J. Agric. Eng. Res. 75, 27–38. doi:10.1006/jaer.1999.0482.

Xu, Z., Yin, J., Ding, J., 2019. Review on control methods and strategies of greenhouse microclimate. DEStech Trans. Comput. Sci. Eng. 0. doi:10.12783/dtcse/icaic2019/29415.

Yan, M., Liu, Z., Wang, J., 2016. Proximal support vector machine improvement and its application to the environmental monitoring of greenhouse plant growth. DEStech Trans. Mater. Sci. Eng. 0. doi:10.12783/dtmse/icmeat2016/6090.

Yang, H., Liu, Q.-F., Yang, H.-Q., 2019. Deterministic and stochastic modelling of greenhouse microclimate. Syst. Sci. Control Eng. 7, 65–72. doi:10.1080/21642583.2019.1661310.

Yang, Y., Zhao, Y., 2012. Prevailing wind direction forecasting for natural ventilation adjustment in greenhouses based on LE-SVM. Energy procedia. In: 2012 International Conference on Future Energy, Environment, and Materials, vol. 16. pp. 252–258. doi:10.1016/j.egypro.2012.01.042.

Yaofeng, H., Meihui, L., Lijun, C., Xiaohui, Q., Shangfeng, D., 2018. Greenhouse modelling and control based on T-S model. In: IFAC-pap., 6th IFAC Conference on Bio-Robotics BIOROBOTICS 2018, vol. 51. pp. 802–806. doi:10.1016/j.ifacol.2018.08.097.

~~Yin, J., Liu, X., Zhang, M., Li, H., 2018. Photosynthetic rate prediction of tomato plant population based on PSO and GA. In: IFAC-pap., 6th IFAC Conference on Bio-Robotics BIOROBOTICS 2018, vol. 51. pp. 61–66. doi:10.1016/j.ifacol.2018.08.061.~~

Yousefi, M.R., Hasanzadeh, S., Mirinejad, H., Ghasemian, M., 2010. A hybrid neuro-fuzzy approach for greenhouse climate modeling. In: 2010 5th IEEE International Conference Intelligent Systems. Presented at the 2010 5th IEEE International Conference Intelligent Systems (IS). IEEE, London, United Kingdom, pp. 212–217. doi:10.1109/is.2010.5548375.

Yu, H., Chen, Y., Hassan, S.G., Li, D., 2016. Prediction of the temperature in a Chinese solar greenhouse based on LSSVM optimized by improved PSO. Comput. Electron. Agric. 122, 94–102. doi:10.1016/j.compag.2016.01.019.

Yue, Y., Quan, J., Zhao, H., Wang, H., 2018. The prediction of greenhouse temperature and humidity based on LM-RBF network. In: 2018 IEEE International Conference on Mechatronics and Automation (ICMA). Presented at the 2018 IEEE International Conference on Mechatronics and Automation (ICMA). IEEE, Changchun, pp. 1537–1541. doi:10.1109/ICMA.2018.8484456.

Zadeh, L.A., 1962. From circuit theory to system theory. Proc. IRE 50, 856–865. doi:10.1109/JRPROC.1962.288302.

Zhang, S., Guo, Y., Zhao, H., Wang, Y., Chow, D., Fang, Y., 2020. Methodologies of control strategies for improving energy efficiency in agricultural greenhouses. J. Clean. Prod. 274, 122695. doi:10.1016/j.jclepro.2020.122695.

Zhang, X., Wang, H., Zou, Z., Wang, S., 2016. CFD and weighted entropy based simulation and optimisation of Chinese Solar Greenhouse temperature distribution. Biosyst. Eng. 142, 12–26. doi:10.1016/j.biosystemseng.2015.11.006.

Zhao, B., Wang, K., 2010. Study on greenhouse environment neural network model based on PSO algorithm. In: 2010 International Conference on Intelligent Computing and Cognitive Informatics. Presented at the 2010 International Conference on Intelligent Computing and Cognitive Informatics (ICICCI). IEEE, Kuala Lumpur, Malaysia, pp. 187–190. doi:10.1109/icicci.2010.69.

Zhou, W., Li, K., Xue, W., Jiang, H., Mao, H., 2017. Multi-objective optimization of fan-pad system operation for venlo greenhouse using CFD model based data interactive mechanism. In: 2017 36th Chinese Control Conference (CCC). Presented at the 2017 36th Chinese Control Conference (CCC). IEEE, Dalian, China, pp. 3020–3025. doi:10.23919/ChiCC.2017.8027822.

Zou, W., Zhang, B., Yao, F., He, C., 2015. Verification and forecasting of temperature and humidity in solar greenhouse based on improved extreme learning machine algorithm. Trans. Chin. Soc. Agric. Eng. 31, 194–200. doi:10.11975/j.issn.1002-6819.2015.24.029.

Highlights

- A systematic literature review of modeling methods for agricultural greenhouse environment is provided.
- Improvement strategies of greenhouse environment and crop model have been summarized and classified.
- Applications of intelligent optimization algorithms for modeling greenhouse have been analyzed.
- Technical trends and valuable research direction to model greenhouse are discussed.

Appendix A Supplementary data

The following is the Supplementary data to this article:

Multimedia Component 1

Multimedia component 1

alt-text: Multimedia component 1

Queries and Answers

Q1

Query: Please confirm that the provided emails “andy_yangwang@cau.edu.cn, wanghongyang1767@gmail.com” are the correct address for official communication, else provide an alternate e-mail address to replace the existing one, because private e-mail addresses should not be used in articles as the address for communication.

Answer: I confirm the emails are the correct address.

Q2

Query: Have we correctly interpreted the following funding source(s) and country names you cited in your article: China Agricultural University, China?

Answer: Yes

Q3

Query: Please provide the volume number or issue number or page range or article number for the bibliography in Ref(s). McNulty, 2017, Xu et al., 2019, Yan et al., 2016.

Answer: McNulty, 2017: it is a magazine article, only having publication date Nov. 3, 2017, please see https://phys.org/news/2017-11-solar-greenhouses-electricity-crops.html

Xu et al., 2019: it is from a proceeding of 2019 International Conference on Artificial Intelligence and Computing Science (ICAICS 2019), pp. 133-141.

Yan et al., 2016.: it is from a proceeding of 2016 5th International Conference on Materials Engineering for Advanced Technologies (ICMEAT 2016), pp. 74-77.

Q4

Query: Uncited references: This section comprises references that occur in the reference list but not in the body of the text. Please position each reference in the text or, alternatively, delete it. Any reference not dealt with will be retained in this section. Thank you.

Answer: Please delete it. We also delete it in the text.

Q5

Query: Please confirm that given names and surnames have been identified correctly and are presented in the desired order and please carefully verify the spelling of all authors' names.

Answer: Yes

Q6

Query: Your article is registered as a regular item and is being processed for inclusion in a regular issue of the journal. If this is NOT correct and your article belongs to a Special Issue/Collection please contact k.kumar.1@elsevier.com immediately prior to returning your corrections.

Answer: Yes